

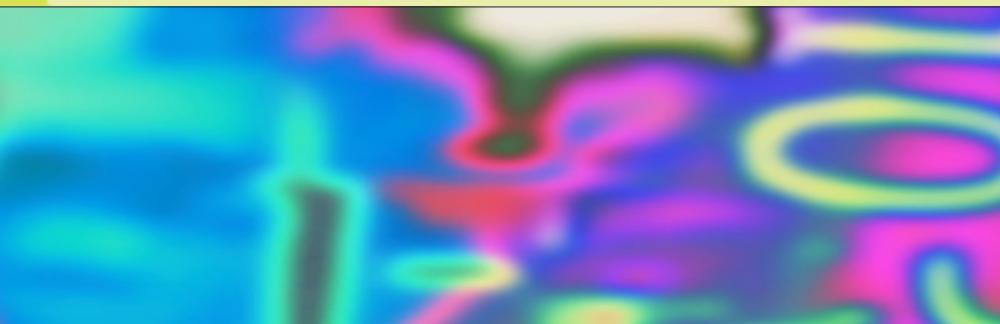
Psicologia Educação e Cultura

Colégio Internato dos Carvalhos

Semestral

Dezembro de 2005

Avaliação Psicológica



Revista de Psicologia, Educação e Cultura

nº 2 Vol. IX

Psicologia Educação Cultura

Revista do Colégio Internato dos Carvalhos

ESTATUTO EDITORIAL

Uma revista semestral e da responsabilidade do Colégio Internato dos Carvalhos e dos departamentos de psicologia, educação e cultura das universidades a que pertencem os membros do Conselho Editorial e do Conselho Consultivo.

Uma revista de carácter científico que pretende acompanhar as diferentes correntes do pensamento acerca da psicologia, da educação e da cultura em geral.

Uma revista que procura actualizar os professores face aos desenvolvimentos recentes na investigação e na prática do ensino-aprendizagem.

Uma revista que pretende capacitar os professores para lidarem com alguns problemas mais frequentes na sala de aula.

Uma revista que vai favorecer a transposição dos estudos no campo da cognição e da afectividade para a prática educativa das escolas.

Uma revista que promove o diálogo entre os professores de diferentes níveis de ensino e possibilita a troca de experiências de sala de aula.

Uma revista que interessa a educadores, professores, investigadores e estudantes, assim como às pessoas que procuram uma formação actualizada, de bom nível, no domínio do ensino-aprendizagem.

CONSELHO CONSULTIVO

Alfonso Barca Lozano (Universidade da Corunha)

Ângela Biaggio (Universidade Federal do Rio Grande do Sul, Brasil)

António Roazzi (Universidade Federal de Pernambuco, Brasil)

Celeste Malpique (Universidade do Porto)

Daniela de Carvalho (Universidade Portucalense, Porto)

David Palenzuela (Universidade de Salamanca)

Etienne Mullet (École Pratique des Hautes Études, Paris)

Feliciano H. Veiga (Universidade de Lisboa)

Francisco C. Carneiro (Universidade do Porto)

Isabel Alarcão (Universidade de Aveiro)

José Tavares (Universidade de Aveiro)

Manuel Ferreira Patrício (Universidade de Évora)

Manuel Viegas Abreu (Universidade de Coimbra)

Maria da Graça Corrêa Jacques (Universidade Federal do Rio Grande do Sul, Brasil)

Nicolau V. Raposo (Universidade de Coimbra)

Paulo Schmitz (Universidade Bona)

Raquel Z. Guzzo (Pontifícia Univ. Católica de Campinas, Brasil)

Rui A. Santiago (Universidade de Aveiro)

Rui Soares (Escola Superior de Educação João de Deus, Lisboa)

Sílvia Koller (Universidade Federal do Rio Grande do Sul, Brasil)

PREÇO E ASSINATURA

Número avulso 10.00 euros

Assinatura/ano 15,00 euros

Inclui os portes dos CTT e IVA à taxa de 5%

SEDE DA REDACÇÃO

Psicologia, Educação e Cultura: Colégio Internato dos Carvalhos

Rua do Padrão, 83 - CARVALHOS

4415-284 PEDROSO

Telefone: 22 786 04 60 Fax: 22 786 04 61

Email: gomes@cic.pt

PROPRIEDADE

P.P.C.M.C.M. - Colégio Internato dos Carvalhos - Cont. N° 500224200

Depósito legal: N° 117618/97

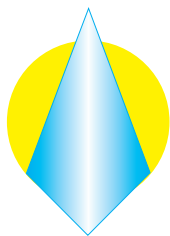
ISSN: 0874-2391

I.C.S.: 121587

N° exemplares: 1000

Capa: anibal couto

P Psicologia E Educação C Cultura



DIRECTOR - EDITOR

João de Freitas Ferreira

SECRETÁRIO

António Fernando Santos Gomes

CONSELHO EDITORIAL

Amâncio C. Pinto (Universidade do Porto)

Félix Neto (Universidade do Porto)

José H. Barros Oliveira (Universidade do Porto)

Leandro S. Almeida (Universidade do Minho)

Joaquim Armando Gomes (Universidade de Coimbra)

Mário R. Simões (Universidade de Coimbra)

Orlando Lourenço (Universidade de Lisboa)

**Os artigos desta Revista estão indexados na base de dados da
PsycINFO, *PsycLIT*, *ClinPSYC* e *Psychological Abstracts*
da American Psychological Association (APA)**

COLÉGIO INTERNATO DOS CARVALHOS

Vol. IX, nº 2, Dezembro de 2005

ÍNDICE

Editorial	
<i>João de Freitas Ferreira e Mário R. Simões</i>	233
Potencialidades e limites do uso de instrumentos no processo de avaliação psicológica	
<i>Mário R. Simões</i>	237
Visualización Espacial y Rendimiento Académico en Dibujo Técnico	
<i>Gerardo Prieto, Rosario Arias-Barahona e Ana-María Núñez</i>	265
Aptidões cognitivas e rendimento académico: A validade preditiva dos testes de inteligência	
<i>Leandro S. Almeida, Gina Lemos</i>	277
Validade preditiva da WPPSI-R em relação ao desempenho escolar ao longo do 1.º Ciclo	
<i>Maria João Seabra Santos</i>	291
A avaliação da inteligência nas dificuldades de aprendizagem: Investigação com a WISC-III	
<i>Marcelino Pereira, Mário R. Simões</i>	307
Inteligência, Educação e Sucesso: As abordagens diferencial e sistémica da Inteligência Humana e suas implicações para a Educação	
<i>Maria João Afonso</i>	337
Competência social: A sua avaliação em contextos de desenvolvimento e educação	
<i>Adelinda Araújo Candeias, Leandro S. Almeida</i>	359
Avaliação da carreira em contexto escolar	
<i>José Tomás da Silva</i>	379
Estudo normativo do teste de Rorschach na população adulta portuguesa: Aspectos metodológicos e alguns resultados	
<i>António Abel Pires</i>	401

Nomeação de Faces Famosas: Capacidade de evocação de nomes próprios numa amostra populacional Portuguesa	
<i>Isabel Pavão Martins, Clara Loureiro, Susana Rodrigues, Beatriz Dias</i>	421
Auto-avaliação da memória do quotidiano: Algumas questões relativas ao uso de questionários de metamemória e estudo das propriedades psicométricas do Questionário de Funcionamento da Memória (QFM).	
<i>Maria Salomé Pinho, Mário R. Simões</i>	437
Motivação intrínseca/extrínseca da vivência religiosa: Uma escala revisitada	
<i>José H. Barros de Oliveira</i>	453
<i>We need more than self-reports...</i> Contributo para a reflexão sobre as estratégias de avaliação do bem-estar	
<i>Rosa Ferreira Novo</i>	477
Pode um grupo pedir perdão a outro grupo? Uma perspectiva timorense	
<i>Félix Neto, Maria da Conceição Pinto, Etienne Mullet</i>	497
Avaliação da credibilidade de alegações de abuso sexual: Consensos e controvérsias	
<i>Carla Machado</i>	513
Construção e avaliação das qualidades psicométricas do Questionário de Gestão do Conhecimento (GC)	
<i>Leonor Cardoso, A. Duarte Gomes, Teresa Rebelo</i>	535
Recensões	557

EDITORIAL

AVALIAÇÃO PSICOLÓGICA

João de Freitas Ferreira e Mário R. Simões

A Avaliação Psicológica constitui uma área que conta já com algum desenvolvimento sistemático em Portugal e é objecto de um interesse crescente por parte de psicólogos, estudantes, professores/educadores, pais, médicos (nomeadamente, pediatras, psiquiatras, neurologistas), juizes e advogados.

Foi preocupação do organizador (Mário R. Simões) do presente número temático da Revista PSICOLOGIA, EDUCAÇÃO E CULTURA, dar conta do desenvolvimento da Avaliação Psicológica em Portugal, bem como de contributos diversificados que surgem em vários contextos do exercício da Psicologia. Houve igualmente a intenção de contar com a colaboração de autores que leccionam e fazem investigação em Avaliação Psicológica nas universidades portuguesas. Neste sentido, este número inclui trabalhos de natureza teórica e investigações de natureza empírica, que abrangem contextos educativos/escolares, clínicos, organizacionais e forenses (procurando assim traduzir a diversidade existente no domínio da Avaliação Psicológica). Conta com participação de docentes e investigadores de universidades nacionais e estrangeiras.

Centrando-nos nos trabalhos apresentados para publicação e procurando salientar alguns elementos possíveis de enquadramento da sua leitura podemos começar por sublinhar que a maior parte dos artigos publicados remete para o uso de instrumentos de avaliação psicológica. Não existe avaliação psicológica formal sem suportes metodológicos ou instrumentais. Daí, a potencial relevância deste tipo de trabalhos.

O primeiro trabalho de Mário R. Simões corresponde justamente a um estudo de natureza teórica (*"Potencialidades e limites do uso de instrumentos no processo de avaliação psicológica"*). Estudos de natureza empírica com instrumentos de avaliação são apresentados por Maria João Seabra Santos, da Universidade de Coimbra (*"Validade preditiva da WPPSI-R em relação ao*

desempenho escolar ao longo do 1.º ciclo"); Marcelino Pereira e Mário R. Simões, da Universidade de Coimbra (*"A avaliação da inteligência nas dificuldades de aprendizagem: Investigação com a WISC-III"*); Isabel Pavão Martins, Clara Loureiro, Susana Rodrigues e Beatriz Dias, da Faculdade de Medicina de Lisboa (*"Nomeação de Faces Famosas: Capacidade de evocação de nomes próprios numa amostra populacional Portuguesa"*); Abel Pires, da Universidade do Porto (Rorschach); Maria Salomé Pinho e Mário R. Simões da Universidade de Coimbra (*"Auto-avaliação da memória do quotidiano: Algumas questões relativas ao uso de questionários de metamemória e estudo das propriedades psicométricas do Questionário de Funcionamento da Memória (QFM)"*); José Barros de Oliveira, da Universidade do Porto (*"Motivação intrínseca/extrínseca da vivência religiosa: Uma escala revisitada"*) e Leonor Cardoso, A. Duarte Gomes e Teresa Rebelo (*"Construção e avaliação das qualidades psicométricas do Questionário de Gestão do conhecimento (QGC)"*).

Noutros trabalhos existe ainda uma referência explícita ao uso de instrumentos específicos de avaliação. É o caso do artigo de Leandro S. Almeida e Gina Lemos, da Universidade do Minho (*"Aptidões cognitivas e rendimento académico: A validade preditiva dos testes de inteligência"*), que analisa a Bateria de Provas de Raciocínio (BPR5/6); do estudo de Adelinda Araújo Candeias, da Universidade de Évora e Leandro S. Almeida, da Universidade do Minho (*"Competências sociais: A sua avaliação em contextos de desenvolvimento e educação"*), que incide sobre uma escala de avaliação de competências sociais; da pesquisa de Rosa Ferreira Novo, da Universidade de Lisboa (*"We need more than self-reports ... Contributo para a reflexão sobre as estratégias de avaliação do bem-estar"*), que recorre às Escalas de Bem-Estar Psicológico (EBEP) e ao Rorschach; do texto de Maria João Afonso, igualmente da Universidade de Lisboa (*"Inteligência, Educação e Sucesso: As abordagens diferencial e sistémica da inteligência humana e suas implicações para a educação"*), que analisa dados do Teste Triárquico de Aptidões de Sternberg – Revisto; da investigação de Gerardo Prieto, Rosario Arias-Barahona, da Universidad de Salamanca e Ana-María Núñez do Instituto de Psicología e Educação do Peru (*"Visualización espacial y rendimiento académico en dibujo técnico"*), que usou o Teste de Visualização (TVZ2002-C); do artigo de Félix Neto e Maria da Conceição Pinto, da Universidade do Porto e Etienne Mullet, da Ecole Pratique des Hautes Etudes de Paris (*"Pode um*

grupo pedir perdão a outro grupo: Uma perspectiva timorense") que recorreu a um questionário de avaliação de atitudes relativas ao "pedir do perdão".

O estudo de José Tomás da Silva, da Universidade de Coimbra (*"Avaliação de carreira em contexto escolar"*), procede igualmente a uma revisão dos principais instrumentos usados na avaliação de carreira. Finalmente, o artigo de Carla Machado, da Universidade do Minho, analisa um tema muito actual em psicologia forense, a *"Avaliação da credibilidade de alegações de abuso sexual: Consensos e controvérsias"*.

Neste número é possível constatar a existência de um conjunto importante de trabalhos que incidem na avaliação de grupos observados em contexto escolar/educativo (crianças em idade pré-escolar, crianças e adolescentes com dificuldades de aprendizagem, estudantes universitários).

Os domínios da inteligência e das aptidões estão igualmente representados em várias investigações (por exemplo, ao nível de instrumentos específicos como a WISC-III ou a WPPSI-R, na análise da relação entre rendimento escolar e aptidões cognitivas através da BPR 5/6 ou do Teste de Visualização, no recurso a diferentes modelos ou teorias da inteligência). Do mesmo modo, a área da avaliação da personalidade é examinada em dois trabalhos com o teste de Rorschach. Outros constructos são igualmente objecto de investigação: a memória (memória do quotidiano, metamemória, memória de nomes de pessoas famosas); o "bem-estar"; o "perdão" e a "motivação intrínseca/extrínseca da vivência religiosa".

De forma esquemática convém ainda sublinhar a diversidade, quer ao nível do tipo de instrumentos usados nos artigos agora publicados (testes de inteligência, testes de aptidões, testes de funções, testes projectivos, questionários e escalas), quer das faixas etárias utilizadas (crianças em idade pré-escolar e escolar, adolescentes, jovens adultos, adultos e idosos).

Finalmente, é de referir de novo o contributo destes trabalhos para a caracterização de práticas de avaliação, em diferentes contextos, no nosso País: escolar, clínico, organizacional e forense.

Resta-nos agradecer a colaboração dos autores dos textos agora publicados e esperar que estes sejam úteis para o ensino, a investigação e a prática da Avaliação Psicológica.

POTENCIALIDADES E LIMITES DO USO DE INSTRUMENTOS NO PROCESSO DE AVALIAÇÃO PSICOLÓGICA

Mário R. Simões*

Serviço de Avaliação Psicológica

Faculdade de Psicologia e de Ciências da Educação,

Universidade de Coimbra, Portugal

Resumo

Caracterizar o papel dos instrumentos de medida em Avaliação Psicológica é o objectivo do presente trabalho. De modo particular são desenvolvidos tópicos relativos à enumeração dos instrumentos de avaliação mais usados, delimitação entre conceitos (teste, *testing* e avaliação psicológica), apresentação de investigações que exemplificam a importância e uso de instrumentos em contextos de avaliação, enumeração de novos desenvolvimentos ao nível da adaptação/construção de instrumentos e das suas normas de utilização, análise de questões de natureza técnica e ética, problemas relativos à interpretação de resultados.

PALAVRAS-CHAVE: *Avaliação Psicológica, Instrumentos de avaliação, Validade, Ética, Interpretação, Investigação.*

Introdução

A *avaliação* e o recurso a *testes* e outros *instrumentos de medida psicológica* constituem um elemento essencial da identidade profissional dos psicólogos e, não menos importante, um aspecto distintivo da especialização em psicologia. Mais especificamente, a avaliação psicológica é parte integrante da prática profissional da psicologia em diferentes **contextos**: clínico (Bénony, 2002; Emmanuelli, 2004; Reynolds, 1998); saúde (Buela-Casal et al., 1996; Cushman e Scherer, 1995; Rozensky, Sweet e Tovian, 2004; Ribeiro, 1999);

Morada (address): Mário R. Simões. Serviço de Avaliação Psicológica. Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra, Rua do Colégio Novo, 3001-802 Coimbra, Portugal. Tel.: 239851450; Fax: 239851462; e-mail: simoesmr@fpe.uc.pt.

* Este texto teve o apoio dos seguintes Projectos de Investigação: "Adaptação e estandardização Portuguesa de Testes Neuropsicológicos: Estudos Normativos e de Validade" [subsidiado pela Fundação para a Ciência e a Tecnologia, no âmbito do Programa POCTI/FEDER (SAPIENS/FCT/35410/PSI/2000) e pelo Centro de Psicopedagogia da Universidade de Coimbra] e "Validação de provas de memória e inventários de avaliação funcional e da qualidade de vida" [financiado pela Fundação Calouste Gulbenkian].

forense (Ackerman, 1999; Gómez, 2001/2003; Heilbrun, 2001, 2003; Pinar e Pagani, 2001; Rogers e Schuman, 2000; Hollin, 2000); educativo (Andrews et al., 2001; Leitão, 2004; Shapiro e Kratochwill, 2000) e organizacional (Edenborough, 2005; Thomas, 2003).

Os instrumentos de medida ocupam em avaliação psicológica um lugar importante, frequentemente decisivo, mas do nosso ponto de vista insuficientemente conceptualizado. Beaune (2004) defende que a *ideia de instrumento* está associada ao verbo construir e reenvia para uma “estrutura material e mental”. Neste plano, pode considerar-se que a “estrutura material” remete para os “materiais” da prova (estímulos), manuais e outras publicações com estudos de natureza teórica e empírica, enquanto a “estrutura mental” considera as estratégias, lógicas e processos de natureza interpretativa subjacentes ao uso de cada prova.

Parece-nos necessária uma reflexão acerca das condições de uso dos instrumentos de avaliação. Nesta medida, pretende-se com o presente trabalho pensar em termos programáticos a função dos instrumentos na avaliação psicológica. Uma tal tarefa inclui:

- (i) a delimitação dos conceitos de “teste”, “testing” e “avaliação psicológica” e uma breve referência ao tipo de instrumentos disponíveis;
- (ii) a análise sintética de práticas actuais representativas da utilização de instrumentos;
- (iii) a identificação dos momentos e circunstâncias da presença do uso de instrumentos;
- (iv) a referência a questões de natureza técnica (regras de construção, procedimentos de utilização) e a exigências de natureza ética e deontológica;
- (v) a definição de uma “biblioteca básica” dos instrumentos de avaliação psicológica.

Neste contexto, o presente estudo actualiza e aprofunda trabalhos anteriores (Simões, 1994; Simões, Almeida e Gonçalves, 1999).

1. Delimitação da natureza dos instrumentos, teste, testing e avaliação psicológica

1.1. São diversificados os tipos de instrumentos que é possível usar em avaliação psicológica. Uma delimitação possível inclui (i) os *inventários de sintomas*, que tal como as *entrevistas estruturadas* estão orientados para a obtenção de descrições de sintomas experimentados (ou para a compreensão de si próprio)

e, nessa medida, dependem da motivação para os comunicar com sinceridade e da capacidade do sujeito para formular juízos exactos; (ii) os *testes de personalidade*, que proporcionam informação acerca de características e estrutura da personalidade; (iii) os *testes de inteligência*, que examinam a capacidade de raciocínio e resolução de problemas, por exemplo; (iv) as *escalas de avaliação*, que facilitam a identificação das percepções dos informadores acerca do comportamento do sujeito; (v) os *procedimentos de observação* e de *auto-monitorização*, que podem envolver, respectivamente, o registo do comportamento operacionalmente definido por parte de observadores independentes (com treino relativamente à técnica e à grelha de registo) ou por parte do próprio sujeito.

1.2. A *avaliação psicológica* é frequentemente conceptualizada como tarefa que remete para a aplicação de testes psicológicos, instrumentos e técnicas (como a entrevista ou a observação directa) (cf. a título de exemplo, Cates, 1999). Neste sentido, é necessário clarificar, desde já, a distinção entre teste, *testing* e avaliação psicológica a partir dos contributos de vários autores (Beutler e Rosner, 1995; Matarazzo, 1990; Fernández-Ballesteros et al., 2001).

O “teste” é um instrumento de medida standardizado que tem como objectivo a obtenção de dados acerca do sujeito. O conceito de “testing” refere-se ao processo delimitado pela administração, cotação e, eventualmente, interpretação dos desempenhos ou respostas individuais a um instrumento (teste, por exemplo). Este processo está associado à atribuição de um significado descritivo às pontuações relativas aos desempenhos, tomando como ponto de referência dados normativos. A aplicação de um teste constitui, de um modo geral, uma tarefa relativamente simples, que pode ser realizada por um psicólogo ou, no limite, por um técnico, eventualmente através do recurso a um computador, e que requer sempre algum tipo de treino ou formação quer genérica, no domínio da utilização de instrumentos de medida standardizados, quer específica, relativa a uma prova particular. Alguns instrumentos usados em avaliação psicológica exigem uma formação aprofundada no que diz respeito à administração e, sobretudo, à cotação e interpretação. É o caso por exemplo de escalas de inteligência (como as escalas de Wechsler), dos inventários de personalidade (como o MMPI ou o Millon) ou de provas projectivas (como o Rorschach). Também por este motivo, consideramos discutível o ponto de vista que defende, em avaliação psicológica, a possibilidade de uma prova ser administrada por técnico com preparação mínima. Na nossa perspectiva a aplicação de um instrumento é indissociável tanto da sua cotação e interpretação, como da integração dos resultados num contexto mais alargado de informações que tem em vista responder ao pedido de avaliação.

A “avaliação psicológica” corresponde quer a um conceito mais amplo, quer a um processo mais exigente onde a ênfase é colocada não apenas nos resultados ou desempenhos específicos resultantes da aplicação de instrumentos, mas também na pessoa que está a ser avaliada, no seu contexto de vida e na resposta ao pedido de exame. Esta actividade inclui o recurso a vários instrumentos e procedimentos de avaliação, que não apenas os testes, e exige um conjunto numeroso de conhecimentos básicos em diferentes domínios da psicologia, isto é, um nível mais elevado de formação e competência. O recurso a instrumentos é apenas uma das tarefas da avaliação psicológica habitualmente presente nas fases inicial (de aquisição de informação) e final (medida da eficácia da intervenção). A escolha de um instrumento depende da natureza do problema a avaliar e nem todas as situações de avaliação psicológica requerem, necessariamente, o uso formal de instrumentos. Por outro lado, a acumulação de dados ou a sua descrição não é “avaliação”. Neste sentido, a avaliação psicológica remete para um processo complexo de tomada de decisão que inclui várias tarefas, operações e procedimentos. Um cuidado especial incide nas *tarefas de inferência e formulação de juízos clínicos*, na *integração* contextualizada dos resultados identificados a partir do recurso a vários instrumentos (testes e outras provas psicológicas), técnicas (entrevista, por exemplo) e informações adicionais comunicadas pelas pessoas que solicitaram a avaliação e provenientes da história e observação do sujeito. A contextualização dos resultados identificados com base nos instrumentos constitui uma tarefa essencial, uma vez que estes são influenciados pelo nível pré-mórbido de funcionamento cognitivo, cooperação, atenção, fadiga, motivação, etc.. A elaboração do relatório psicológico exemplifica bem a complexidade associada ao trabalho de avaliação psicológica.

1.3. Num registo complementar, vários investigadores (Krishnamurthy et al., 2004; Meyer et al., 2001) definem a actividade de avaliação psicológica como um processo complexo, intensivo e multifacetado que exige: (i) a compreensão da personalidade, envolvendo conhecimentos quer da psicopatologia quer relativos ao modo como as perturbações neurológicas se manifestam no funcionamento cognitivo e no comportamento; (ii) o domínio das teorias e modelos de avaliação, diagnóstico e medida psicológica; (iii) o conhecimento da psicometria e dos métodos de investigação e construção de instrumentos; (iv) a competência de aquisição sistemática de dados, selecção e uso (administração, cotação e interpretação) de métodos, técnicas e instrumentos de avaliação (e diagnóstico) das dimensões cognitiva, afectiva e comportamental da experiência humana; (v) o reconhecimento que diferen-

tes métodos e instrumentos de avaliação têm vantagens e limitações específicas e fornecem tipos qualitativamente distintos de informação; (vi) a capacidade para conceptualizar condições de vida real que podem dar origem a um padrão particular de dados nos testes; (vii) o saber desafiar o próprio juízo associando, sistematicamente, a presença e ausência de indicadores de teste às características psicológicas examinadas; (viii) a competência interpessoal e a sensibilidade emocional para comunicar eficazmente com as pessoas avaliadas, outras pessoas significativas, fontes de encaminhamento dos casos; (ix) as competências de integração da informação, análise e inferência; (x) a capacidade de conceptualização ou formulação psicológica do caso e a elaboração de relatórios; (xi) as competências técnicas de identificação do problema e definição de objectivos e recomendações aceitáveis de intervenção; (xii) a capacidade para compreender as relações entre avaliação, planificação e medida dos resultados da intervenção, o que implica saber monitorizar o progresso decorrente da intervenção e medir os seus resultados. Nesta caracterização do processo de avaliação psicológica está contemplado, em vários momentos, o recurso explícito ou implícito de testes e outros instrumentos.

Por outro lado, a avaliação psicológica formal pode ser prolongada no tempo quando é implementada antes, durante e após a intervenção (aconselhamento, psicoterapia, reabilitação): em primeiro lugar, para compreender e identificar o problema do sujeito e, caso necessário, definir o diagnóstico (Barron, 1998; Beutler e Malik, 2002); depois, para organizar a intervenção (Hersen e Porzelius, 2002) e, por último, para medir a eficácia dos resultados (Simões e Lopes, 2004). A avaliação focalizada exclusivamente no diagnóstico nosográfico corresponde actualmente a um pequeno número de pedidos de consulta.

2. Importância do uso de instrumentos em avaliação (neuro)psicológica: Resultados de algumas investigações

A importância da tarefa de avaliação psicológica pode ser examinada a partir de um conjunto de investigações realizadas com psicólogos. Estas pesquisas incidem sobre a presença ou exigência da avaliação e dos respectivos instrumentos em contextos de ensino da psicologia pré e pós-graduado e na prática profissional, por exemplo. Os resultados mostram que: (i) a seguir à psicoterapia, a avaliação psicológica constitui a actividade clínica mais importante dos psicólogos (Meyer et al., 1998); (ii) 91% dos psicólogos estão envol-

vidos nalguma forma de avaliação (Watkins et al., 1995); (iii) 64% dos anúncios de pedidos de psicólogos listam a competência em avaliação como um importante pré-requisito para preenchimento do lugar (Kinder, 1994).

Ao mesmo tempo, a *competência específica* em avaliação psicológica e no uso de testes e outros instrumentos em particular constituem pré-requisitos importantes na admissão a "internatos pré-doutoramento" (Stedman, Hatch e Schoenfeld, 2000) e elementos essenciais dos *curricula* de programas de doutoramento (Piotrowski e Zalewski, 1993). Nestes últimos, assiste-se, de um modo geral, à manutenção ou aumento da ênfase na componente de avaliação psicológica (Piotrowski e Belter, 1999). A sugestão, mais rara, de uma ligeira diminuição é atribuída ao declínio no ensino das técnicas projectivas (Belter e Piotrowski, 2001).

O aparente decréscimo na utilização intensiva de alguns instrumentos, que historicamente definiram a prática profissional dos psicólogos e a diminuição do tempo dedicado a tarefas de avaliação (Piotrowski, 1999), que passou de 50% (Sundberg, 1961) para 22% em 1982 (Lubin et al., 1984, e 20% em 1995 (Watkins et al., 1995), sugerido nalguns estudos, parece resultar principalmente das restrições ao reembolso destas actividades por parte de companhias de seguro e outras instituições que financiam este tipo de prestação de serviços por parte de psicólogos (Eisman et al., 2000) e a um menor uso das técnicas projectivas.

Todavia, existem investigações que, em contextos de avaliação neuropsicológica, contrariam fortemente esta sugestão. Ainda pouco desenvolvida em Portugal, a avaliação neuropsicológica constitui uma área particular da avaliação psicológica onde é possível observar um crescimento acentuado no uso e investigação clínica com testes. Neste âmbito, Camara e cols. (2000) pesquisaram o uso de testes por parte dos psicólogos clínicos filiados na *American Psychological Association* (APA) e na *National Academy of Neuropsychology* (NAN) e constataram que mais de 80% dos psicólogos clínicos ocupavam menos de 5 horas por semana em actividades de avaliação. Um padrão totalmente distinto foi comunicado por parte dos neuropsicólogos da NAN: 35% dos neuropsicólogos trabalhava entre 10 a 20 horas por semana em avaliação e 33% mais de 20 horas. Por sua vez, numa pesquisa conduzida junto de membros da Divisão 40 da APA (Neuropsicologia Clínica) e da NAN, Sweet et al. (2002) verificaram que o tempo da avaliação neuropsicológica era predominantemente utilizado em tarefas de aplicação de testes, seguido de actividades de interpretação e elaboração de relatórios, cotação de testes e entrevista clínica. Mais recentemente, Rabin et al. (2005) concluíram, numa amostra de psicólogos filiados na APA (Divisão 40), na NAN e na *Internatio-*

nal *Neuropsychological Society* (INS), que a percentagem de tempo dedicada à avaliação neuropsicológica (42%) é substancialmente mais elevada do que a empregue na realização de outras tarefas: psicoterapia (19.2%), investigação e/ou ensino (16.7%), reabilitação e/ou remediação cognitiva (8%) e outras actividades (13.9%). Vale ainda a pena referir que as aptidões ou funções objecto de avaliação frequente, quase sempre examinadas através do recurso exclusivo a instrumentos (testes), incluíam a atenção (97.7%), memória verbal (96.1%), funções executivas (96.3%), competências viso-espaciais (94.6%), memória não verbal (87.8%), inteligência (84.9%), linguagem (83.5%), capacidades construtivas (83.7%), competências motoras (67.9%), rendimento escolar (54.9%), percepção auditiva (43.5%), percepção táctil (27.2%). Nesta pesquisa, os *pedidos de avaliação*, que implicavam frequentemente o recurso a testes, englobavam a definição do diagnóstico (70.7%), a reabilitação/planificação da intervenção (48.3%), a decisão forense (31.8%), a planificação educativa (29.6%), a avaliação da capacidade para trabalhar (27.9%), a definição da linha-base de uma função para posterior exame (24.3%), a avaliação da capacidade para viver autonomamente (19.7%), o exame de intervenções médicas (9.6%), outro pedido (4.3%) e a localização da lesão (2.7%).

Para concluir, convém adicionar a esta listagem de trabalhos outros estudos (cf. Archer, et al., 1991; Bartram e Coyne, 1998; Camara et al, 2000; Lees-Haley et al, 1996; Muñiz et al, 1999, 2001; Piotrowski e Lubin, 1990), que sugerem o recurso em grande escala e a utilidade e eficácia de instrumentos em avaliação psicológica.

3. Crescente especificidade e complexidade técnica no uso de instrumentos de avaliação

Tem-se assistido nos últimos anos a importantes desenvolvimentos de natureza programática ("normas", "regras", "princípios") que são essenciais à regulação do exercício profissional da avaliação e à definição das "melhores práticas" no uso de instrumentos de avaliação. Estes desenvolvimentos traduzem uma progressiva consciência dos usos apropriados e limitações no recurso a instrumentos de avaliação e ensaiam respostas para problemas conhecidos relativos à construção e adaptação de instrumentos. Mais especificamente, estes "novos desenvolvimentos" abrangem: (i) a definição de linhas orientadoras para a avaliação dos instrumentos de avaliação standardizados (Cichetti, 1994); (ii) normas para a construção de testes que têm subjacente uma cres-

cente especialização e exigência técnicas (*American Psychological Association*, 1999; Downing e Haladyna, no prelo; Eignor, 2001; Embretson e Reise, 2000; Fredericksen et al., 1993; Irvine e Kyllonen, 2001; Urbina, 2004; Thissen e Wainer, 2001; Wilson, 2005); (iii) uma reflexão centrada nos constructos (Braun, Jackson e Wiley, 2002); (iv) métodos e práticas de adaptação de testes a diferentes culturas (Geisinger, 1994; Hambleton, 2001; Hambleton, Merenda e Spielberger, 2005); (v) o uso de instrumentos de avaliação em contextos multi-culturais (Dana, 2005; Gopaul-McNicol e Armour-Thomas, 2002; Suzuki, Ponterotto e Meller, 2001); (vi) novas tecnologias associadas à avaliação computadorizada (Drasgow e Olson-Buchanan, 1999; Maarse et al., 2003; Wainer, 2000); (vii) problemas levantados pela colocação de testes na Internet (Naglieri et al., 2004); (viii) identificação das competências ou qualificações necessárias ao uso de testes (Bartram, 2001; Moreland et al., 1995; Turner et al., 2001); (ix) uso responsável dos testes (*American Psychological Association*, 2002; Eyde et al., 1993); (x) enumeração dos obstáculos à utilização de testes (Eisman et al., 2000; Meyer et al., 1998); (xi) caracterização de problemas de natureza ética associados ao uso de instrumentos (*American Psychological Association*, 2002).

Dois destes tópicos merecem breves comentários: as questões de natureza técnica, por um lado e os problemas de natureza ética, por outro.

4. Instrumentos de avaliação: questões de natureza técnica

O processo de validação e aperfeiçoamento dos instrumentos remete para tarefas de âmbito técnico e corresponde, igualmente, a exigências de natureza ética.

4.1. Deve reconhecer-se que uma maior fiabilidade das aproximações ao funcionamento psicológico do sujeito depende da existência de informações e resultados convergentes, obtidos ao longo do tempo, a partir do recurso a várias fontes: pessoas, registos, instrumentos. Sabemos que o uso de vários tipos de instrumentos (abordagem multi-método) potencia a obtenção de uma amostra mais representativa do comportamento e contraria parte dos limites inerentes à utilização de cada prova em particular. Neste sentido, considerar os sintomas e a dinâmica dos comportamentos a partir de diversas perspectivas, introduzidas através do recurso a diferentes instrumentos (e interlocutores), constitui uma forma de assegurar uma maior objectividade ao trabalho de avaliação e de captar as variações no comportamento e na experiência do sujeito.

Do ponto de vista da técnica é necessário considerar os seguintes aspectos. Em primeiro lugar, é essencial aperfeiçoar o “modelo” relativo ao uso de cada instrumento de avaliação. Neste plano, o respectivo manual deve: (i) proporcionar informação específica e uma representação rigorosa do constructo que pretende medir, tentando contrariar a sua parcialidade ou incompletude; assim, são desejáveis discriminações finas do constructo que se pretende avaliar com os resultados; (ii) incluir resultados de investigações que refiram elementos relativos a rubricas como a análise dos itens, a fiabilidade (consistência interna, estabilidade temporal, erro de medida, acordo entre codadores, por exemplo) e a validade (conteúdo, concorrente, preditiva, constructo, convergente e discriminante, por exemplo). No âmbito da validade, são essenciais os resultados de pesquisas realizadas quer com grupos normativos, quer com grupos especiais relevantes (de natureza clínica, educativa ou forense, consoante os casos). As tarefas de validação de um instrumento nunca estão definitivamente concluídas. Neste plano, é imperioso que o utilizador da prova esteja atento e actualizado em relação aos resultados provenientes de novos estudos que corroborem ou infirmem o valor da prova. Por outras palavras, o manual da prova é apenas um ponto de partida essencial.

Em segundo lugar, é forçoso assumir que os resultados num único instrumento correspondem a indicações, são sempre aproximações à realidade do funcionamento psicológico do sujeito, que devem ser aperfeiçoadas através da fundamentação empírica das interpretações formuladas a partir do uso de uma bateria de instrumentos de avaliação. É, por isso, necessário considerar o erro de medida associado ao recurso integrado de uma bateria de provas psicológicas e, não apenas, como é comum, o erro relativo ao uso de um instrumento particular.

Em terceiro lugar, a fiabilidade e validade dos instrumentos habitualmente examinadas a nível de grupo devem ser analisadas, também, a um nível individual. Esta nova estratégia, orientada para a validação dos instrumentos considerados individual e agregadamente, por um lado, e ao nível do grupo e de cada sujeito particular, por outro, corresponde a uma forma mais exigente de aferir o valor e utilidade dos instrumentos.

4.2. O *aperfeiçoamento de um instrumento* refere-se a um processo contínuo de natureza conceptual e técnica que pode conduzir a avanços na teoria, na identificação de novos constructos ou na modificação das definições do constructo, tornando-os mais precisos. Este refinamento remete para um conjunto de procedimentos designados para melhorar a representação do constructo medido e que se aplicam quer à construção de novos instrumentos (cf.

Clark e Watson, 1995), quer ao aperfeiçoamento dos já existentes, incluindo a sua adaptação a novos usos com diferentes populações. Para isso, é necessário especificar, de modo preciso, o constructo medido e assegurar que o instrumento representa com exactidão esse constructo (Foster e Cone, 1995). Neste âmbito, são necessárias várias análises complementares tendo em vista: (i) a identificação da estrutura da prova (número e natureza dos factores ou dimensões presentes especificados através de técnicas de análise factorial exploratória e/ou confirmatória; esta última técnica permite a avaliação do grau de ajustamento de um modelo aos dados e a comparação sistemática entre modelos alternativos); (ii) o estudo da consistência interna (caracterizada através do coeficiente alfa de Cronbach, das correlações item-total ou entre itens); o seu aperfeiçoamento pode passar pela introdução de novos itens, revisão ou eliminação de itens já existentes; (iii) a análise da homogeneidade do conteúdo de cada dimensão; (iv) a inclusão de itens que discriminem desempenhos ou respostas dos sujeitos a um nível desejado de intensidade do atributo; pretende-se assim contar com itens que diferenciem por exemplo sujeitos de grupos clínicos de sujeitos da população normal; (v) a confirmação das propriedades psicométricas do instrumento em amostras independentes; (vi) o exame da validade discriminante, que permita diferenciar um constructo de outro e o desenvolvimento de medidas suficientemente específicas (Smith e McCarthy, 1995).

4.3. Um desenvolvimento relativamente recente no domínio da construção de instrumentos remete para a Teoria de Resposta aos Itens (TRI). Trata-se de uma abordagem que procura ultrapassar as limitações dos métodos e modelos associados à Teoria Clássica dos Testes (TCT). No que diz respeito à estatística dos itens (nível de dificuldade e nível de discriminação) ela depende na TCT das amostras utilizadas e, por isso, não pode ser necessariamente generalizada para a população (a TRI assegura invariância da estatística dos itens independentemente das amostras). Por outro lado, na TCT, a comparabilidade relativa aos resultados é limitada às situações em que os sujeitos responderam ao mesmo instrumento (na TRI a estimativa da “aptidão” (*ability*) ou de qualquer outro constructo medido não está dependente do instrumento usado). Em termos muito resumidos, a TRI envolve dois postulados: (i) subjacente ao desempenho do sujeito existe uma única dimensão, uma aptidão ou um traço de personalidade, por exemplo; (ii) a relação entre a probabilidade do sujeito apresentar uma resposta correcta e a “aptidão” do sujeito pode ser descrita a partir de uma curva (“curva característica do item”) que representa a relação entre a probabilidade de resposta correcta e a “aptidão”.

4.4. Um outro desenvolvimento remete para a crescente disponibilização de *software* para a avaliação psicológica computadorizada. A avaliação apoiada no recurso a computadores (e a avaliação *online*, através da Internet) assegura serviços de acesso mais rápido, com custos económicos mais reduzidos e uma maior uniformização e consistência na administração e cotação em comparação com a avaliação realizada em situação de face-a-face. O interesse por esta forma de avaliação (é cada vez maior o número de testes adaptados para suportes informáticos) não tem tido, no entanto, correspondência ao nível dos programas de formação (Olson, 2001). No caso da avaliação *online*, é necessário reconhecer a existência de problemas com a autenticidade dos dados (confirmação da identidade da pessoa que responde aos itens), a segurança relativa aos materiais de teste, a natureza das normas a usar e a interpretação dos dados (Barak e English, 2002; Buchanan, 2002; Butcher et al., 2000; Butcher et al., 2004; Naglieri et al., 2004). Por outro lado, aquilo que está disponível nos suportes informáticos (e na Internet) são testes ou instrumentos de avaliação e não a avaliação psicológica propriamente dita. Esta requer uma relação de confiança que não é sinónimo de relação impessoal. Além disso, não são os testes (ou o computador) os responsáveis pela interpretação dos resultados, mas sim o psicólogo.

4.5. A construção e a utilização de instrumentos seguem um conjunto numeroso de regras precisas que não se compadece com improvisações ou simplificações cómodas. Uma delas enfatiza a ideia de quantificação da realidade psicológica tornada possível com o uso de instrumentos científicos e remete para a precariedade de usos superficiais que concedem um privilégio absoluto aos resultados dos testes, que traduziriam a verdade acerca do sujeito – uma realidade profunda supostamente descrita de um modo exaustivo e exacto, com uma precisão absoluta. O “real é aquilo que pode ser medido”, escreve Max Planck mas, ao mesmo tempo, assevera Beaune, um “enigma” que ressurge continuamente. Encontramos neste ponto elementos de contacto com o próprio conceito psicométrico de validade, tal como ele é caracterizado por autores como Campbell, Fiske, Cronbach, Meehl, Messick, Shepard, entre outros. Em termos muito esquemáticos, a validade: (i) é um conceito unitário que requer várias fontes de evidência; (ii) corresponde à característica mais importante subjacente ao uso de um instrumento; (iii) remete para os resultados e não para o instrumento; (iv) refere-se ao grau com que a informação proporcionada pelas pontuações num instrumento é útil, apropriada e objectiva; (v) indica em que medida as interpretações (inferências) definidas a partir dos resultados estão fundamentadas do ponto de vista científico (ou seja, o grau em

que as interpretações dos resultados usadas para fins específicos são apoiadas pela teoria e por evidência de natureza empírica); (vi) baseia-se nos dados provenientes de várias investigações. Nesta linha, a validação de um instrumento corresponde a um processo contínuo associado a um conjunto de tarefas complexas que supõem estudos de validade de conteúdo (os quais exigem o envolvimento de peritos na definição e representação do domínio e supõem a representatividade dos comportamentos associados a um constructo particular); pesquisas de validade concorrente (que requerem a análise das correlações entre os resultados no instrumento que se pretende validar e os resultados noutra prova que reconhecidamente examina o mesmo constructo); investigações no âmbito da validade preditiva (apelando para a existência de uma relação substancial entre os desempenhos na prova e os resultados obtidos mais tarde de acordo com um critério externo pertinente); estudos no domínio da validade de constructo (supondo, por exemplo, a obtenção de dados sobre a validade convergente e a validade discriminante da prova ou acerca da sua estrutura interna); (vii) considera as consequências sociais do programa de avaliação subjacente ao uso do instrumento. Por outras palavras, a credibilidade científica da avaliação com um instrumento depende da natureza dos dados disponíveis acerca da sua validade que, no entanto, nunca está definitiva ou absolutamente estabelecida e não é uma “questão de tudo ou nada”.

Como veremos seguidamente, as questões de natureza técnica e ética são tópicos indissociáveis.

5. Instrumentos de avaliação: questões de natureza ética

Clero (2004) sugere que “todo o instrumento abre diante de si um horizonte de perfeição”. Todavia, convém tomar consciência das dificuldades e limites, é necessário reflectir sobre os riscos apesar das potencialidades técnicas, (re)pensar as lógicas de utilização, o que fazemos com os instrumentos. Nesta linha, e do ponto de vista das preocupações éticas podemos colocar algumas questões: quem avalia as práticas de utilização de instrumentos? Que efeitos ou consequências é possível antecipar das suas aplicações? Em que medida a técnica dos testes, uma escala de inteligência por exemplo, “ajuda” a fabricar fenómenos ou provoca artificialmente a sua aparição (“deficiência mental”, “sobredotação”)?

Neste particular, à semelhança do que acontece com qualquer tecnologia científica, os testes ou outros instrumentos de avaliação podem ser usados de

forma apropriada ou incorrectamente. No primeiro caso, a informação obtida pode ser de grande utilidade para compreender o sujeito. No caso do uso inapropriado, a informação veiculada distorce a verdade acerca do sujeito. Neste contexto, os Códigos de Ética e Deontologia introduzem também princípios que orientam a tomada de decisão no domínio do uso de instrumentos. A este propósito, na Secção 9 (Avaliação) do documento *Princípios Éticos dos Psicólogos e Código de Conduta* (American Psychological Association, 2002): (i) considera-se que os psicólogos que desenvolvem testes e outras técnicas de avaliação devem usar procedimentos psicométricos apropriados e conhecimentos científicos e profissionais actualizados no que diz respeito ao desenho, standardização, validação, redução ou eliminação de enviesamento e recomendações para a sua utilização (9.05. Construção de Testes); (ii) reconhece-se o papel dos psicólogos quer na administração, cotação e interpretação apropriada de técnicas de avaliação, entrevistas, testes ou instrumentos tendo em conta os dados da investigação ou a evidência da sua utilidade, quer na escolha dos instrumentos a usar – neste particular, a sua fiabilidade e validade devem estar bem definidas para a população na qual vão ser empregues (9.02. Usos da Avaliação); a aplicação, cotação e interpretação apropriadas dos instrumentos é da responsabilidade dos psicólogos (9.09 “Cotação dos testes e Serviços de Interpretação”), bem como a explicação dos resultados (9.10 “Explicação dos Resultados da Avaliação”) e a protecção dos materiais (manuais, instrumentos, protocolos, questões, estímulos) (9.11. “Preservar a Segurança dos Testes”); (iii) defende-se a necessidade de fundamentação das opiniões contidas nas recomendações, relatórios, afirmações de natureza avaliativa ou diagnóstica com base em informação e técnicas suficientes que legitimem os dados obtidos; as opiniões relativas a características psicológicas apenas devem ser formuladas após a realização de um exame cujos resultados apoiem as referidas conclusões; deve ser considerado o impacto da informação insuficiente na fiabilidade e validade das opiniões, conclusões e recomendações (9.01. “Bases para a Avaliação”); o que significa que as decisões de avaliação ou intervenção devem ser baseadas no recurso a testes actualizados (9.08 “Testes Obsoletos e Resultados Ultrapassados nos Testes”).

Num registo complementar é de contestar algumas práticas de sinal mais conformista: os instrumentos mais frequentemente usados são aqueles que se encontram mais facilmente acessíveis e, de modo particular, os que foram objecto de ensino/aprendizagem ou treino académico (Fischer, 1992; Marlowe et al., 1992; Stout, 1992; Watkins, 1991).

Tal prática contraria vários outros princípios orientadores complementares a ter em conta no recurso a instrumentos de avaliação: (i) adquirir e manter

atualizados conhecimentos e competências relativos ao uso de vários instrumentos disponíveis; (ii) reconhecer o nível de conhecimentos ou formação especializada exigidos no uso de instrumentos específicos (a aprendizagem, formação teórica, investigação e treino prático com provas de avaliação psicológica constitui um aspecto essencial); (iii) saber examinar o valor e utilidade prática dos instrumentos, incluindo aquilo que os resultados podem ou não dizer.

6. Instrumentos de avaliação: observação, conceptualização ou formulação do caso, juízo clínico

Uma “boa avaliação” é aquela que assegura uma definição exacta do retrato ou representação do sujeito. É no entanto necessário inscrever a avaliação (e o recurso a instrumentos) na sua temporalidade. A avaliação é um “instantâneo fotográfico” que descreve um momento fixado no tempo, descrito a partir do ponto de vista do psicólogo (Cates, 1999).

Os instrumentos proporcionam informação quantitativa e qualitativa que viabiliza uma avaliação empírica mais precisa acerca dos traços e outras características. É, no entanto, necessário reconhecer a pertinência das críticas à *ênfase excessiva atribuída frequentemente aos resultados quantitativos* (QI's, resultados padronizados, percentis). Neste sentido deve ser dada mais atenção ao *comportamento durante a situação de teste*. Uma mesma pontuação num teste não tem o mesmo significado quando é alcançada por um sujeito que “manifesta interesse pelas tarefas de teste”, “persistência nas tarefas difíceis, não desiste com facilidade” e “se esforça por resolver os itens do teste” ou, pelo contrário, por um sujeito que “manifesta falta de envolvimento e interesse”, “permanece desatento” ou “necessita de elogios para completar as tarefas”.

Para Hunt (1946) ainda que importante, a fixação nos resultados quantitativos pode obscurecer a presença de informação clinicamente útil e, por isso, a preocupação com as observações qualitativas (e as diversas *nuances* do comportamento do sujeito) deve estar presente quer na construção de testes (construir um teste de modo que este conduza a um máximo de respostas qualitativas), quer no treino clínico dos psicólogos. Aqui, o programa pode ser: explorar o significado das respostas e dos resultados, tentar compreender as razões das pontuações de um sujeito particular (Sweet, Moberg e Westergaard, 1996), usar estratégias para compreender melhor os processos mobilizados pelo sujeito para dar uma resposta parti-

cular (Kaplan, Fein, Morris e Delis, 1991; Groth-Marnat, Gallagher, Hale e Kaplan, 2000).

O recurso a diferentes instrumentos serve para construir e verificar hipóteses. Enquanto dispositivo técnico o instrumento encontra-se associado à produção de saber, torna possível o conhecimento psicológico. Como escreve Barthélémy (2004) “apenas o instrumento pela sua função de conhecimento pode concretizar plenamente a técnica” (p. 50). Neste sentido, os instrumentos possibilitam a optimização do trabalho do psicólogo. No entanto, nenhum resultado ou conjunto de resultados num único instrumento proporciona todas as respostas ou o saber necessário acerca do sujeito. É, por isso, necessário o recurso sistemático a uma bateria de testes como um meio de testar hipóteses. Mas, esta opção por um conjunto alargado de provas é, ainda assim, insuficiente. Neste plano, a competência inferencial do psicólogo, isto é, a capacidade deste para formular juízos a partir da informação disponível obtida a partir do recurso a métodos, técnicas e instrumentos de medida psicológica, constitui uma questão central na determinação da qualidade da avaliação. Este tópico não será aqui não desenvolvido.

A articulação dos dados provenientes dos instrumentos é do domínio do juízo clínico. A inferência e a intuição clínicas desempenham um papel decisivo na interpretação dos dados obtidos e constituem um aspecto importante da conceptualização do caso (tópico essencial da entrevista de restituição da informação e do relatório). Meyer et al. (2001) lembram a este propósito que os testes e outros instrumentos de avaliação “não pensam por si próprios, nem comunicam directamente com os pacientes”. Tal como o estetoscópio, o aparelho de medir a pressão arterial ou a ressonância magnética, “um teste psicológico é um instrumento que não fala, cujo valor não pode ser separado da sofisticação do clínico que desenha as inferências a partir dos resultados e comunica com pacientes e outros profissionais”. Na formulação de Muñiz et al. (2001) é preciso considerar os testes como instrumentos potencialmente úteis “mas não como uma espécie de varinha mágica para resolver problemas e tomar decisões”.

Gaston Bachelard refere que os microscópios foram inventados para alargar a razão e não a visão (Bachelard, 1938/1996). Na perspectiva da avaliação psicológica, o essencial dos instrumentos está no facto deles permitirem compreender melhor, medir, decifrar o que ainda não se sabe e, nessa medida, na possibilidade de através deles se pensar melhor o sujeito. Mas a tarefa de conceptualização ou formulação de caso não acaba com a identificação das pontuações nos instrumentos usados. Ela tem subjacente um ideal de precisão quer do funcionamento psicológico actual, quer das previsões relativas

ao comportamento no futuro. A este respeito o leitor pode apreciar o debate “predição actuarial versus predição clínica” (Garb, 1998; Grove et al., 2000; Hunt, 1946; Meehl, 1954; Waller, Yonce, Grove, Faust e Lenzenweger, 2005).

7. Elementos para uma biblioteca básica dos utilizadores de instrumentos de avaliação psicológica

A tarefa de familiarização com novos instrumentos disponíveis e com as revisões dos já existentes constitui uma exigência elementar ao exercício da avaliação psicológica. Neste plano, convém sublinhar a existência de um número considerável de publicações periódicas e livros (editados na última década) que no seu conjunto constituem parte integrante da biblioteca básica em avaliação psicológica. Para além das *enciclopédias* (Fernandez-Ballesteros, 2003; Graham e Naglieri, 2003), incluem-se aqui um conjunto numeroso de obras centradas na revisão independente de instrumentos específicos de avaliação psicológica/neuropsicológica/psicopatológica (Bouvard, 1999; Bouvard e Cottiaux, 1998; Burns, Lawlor e Craig, 2004; Dorfman e Hersen, 2001; Franzen, 2000; Golden et al., 2000; Grisso, 2003; Gutiérrez et al., 2003; Hodges, 1995; Lezak et al. 2004; López et al., 2002; Lopez e Snyder, 2003; Maruish, 2004a, 2004b; McCaffrey et al., 2000; Mitrushina et al., 2005; Newark, 1996; Shulman e Feinstein, 2003; Spreen e Strauss, 1998; Wade, 1995). O trabalho de Lopez e Snyder (2003) apresenta um interesse particular na medida em que está centrado em constructos ou dimensões positivas do funcionamento psicológico (*locus* de controlo; auto-estima; optimismo aprendido; optimismo; esperança; auto-eficácia; resolução de problemas; criatividade; sabedoria; coragem; emoções positivas; amor romântico; inteligência emocional; empatia; vinculação; perdão; sentido de humor; gratidão; bem-estar subjectivo; qualidade de vida).

Ainda no que diz respeito a livros focalizados em instrumentos, são de assinalar os livros mais recentes da Série *Mental Measurements Yearbooks*, iniciada em 1938 (Plake e Impara, 2001, Vol. XIV; Plake, Impara e Spies, 2003, Vol. XV; Spies e Plake, 2005, Vol. XVI; em cada volume são habitualmente recenseados e revistos habitualmente mais de duas centenas de instrumentos), e da Série *Tests in Print*, iniciada em 1961 (Murphy, Plake, Impara e Spies, 2002, Vol. VI);

Entre nós, são de referir as Séries “Provas Psicológicas em Portugal” (23 instrumentos analisados; Almeida, Simões e Gonçalves, 1995); “Testes e Provas Psicológicas em Portugal” (17 instrumentos revistos; Simões, Gonçal-

ves e Almeida, 1999) e “Avaliação Psicológica: Instrumentos validados para a população portuguesa” (Vol. I: 14 instrumentos recenseados; Gonçalves, Simões, Almeida e Machado, 2003; Vol. II: revisão de 10 instrumentos; Almeida, Simões, Machado e Gonçalves, 2004; Vol. III: 10 instrumentos analisados; Simões, Machado, Gonçalves e Almeida, no prelo). No âmbito dos livros é necessário referir igualmente a colecção “*Essentials of Psychological Assessment Series*”, da editora Wiley, iniciada em 1999, e que conta já com aproximadamente três dezenas de títulos publicados, uma parte importante dos quais centrados em: (i) instrumentos de avaliação da inteligência (WISC-III e WPPSI-R; WISC-IV; WPPSI-III; WAIS-III; CAS; KAIT; *WJ III Cognitive Abilities*; *Stanford-Binet Intelligence Scales/SB5*); (ii) instrumentos de avaliação neuropsicológica (WMS-III; NEPSY); (iii) instrumentos de exame do desenvolvimento (*Bayley Scales of Infant Development-III*); (iv) instrumentos de avaliação da personalidade, incluindo técnicas projectivas (TAT; Rorschach) e inventários de personalidade (Millon; MMPI-2; MMPI-A; PAI; 16PF); (v) instrumentos de avaliação do rendimento ou realização escolar (*WJ III Achievement Tests*).

Por outro lado, a análise das numerosas publicações periódicas na área da avaliação psicológica (*Psychological Assessment*; *Assessment*; *Journal of Personality Assessment*; *European Journal of Psychological Assessment*; *Journal of Psychoeducational Assessment*; *Educational and Psychological Measurement*; *Journal of Psychopathology and Behavioral Assessment*; *International Journal of Testing*; *International Journal of Selection and Assessment*; *Revista Ibero-Americana de Diagnóstico e Avaliação Psicológica*; *Avaliação Psicológica*) mostra bem a relevância atribuída à investigação com instrumentos nesta área. No domínio mais específico da avaliação neuropsicológica (e dos respectivos instrumentos) são de assinalar, nomeadamente, as seguintes revistas: *The Clinical Neuropsychologist*; *Archives of Clinical Neuropsychology* e *Journal of Experimental and Clinical Neuropsychology*.

Em Portugal é de destacar, no que se refere a publicações periódicas, os seguintes números temáticos: *Psicologica* [1994, Vol. 11 (“Avaliação Psicológica: Estudos teóricos”), 1994, Vol. 12 (“Avaliação Psicológica: Estudos empíricos”); 2003, Vol. 34 (“Avaliação Neuropsicológica”)]; *Psicologia* [1989, Vol. VII, n.º 3 (“Psicometria e Avaliação Psicológica”)]; *Psicologia, Educação e Cultura* [2005, Vol. IX, n.º2 (“Avaliação Psicológica”)] e as Actas dos Congressos “Avaliação Psicológica: Formas e Contextos” (1993-1999, 2004, 7 volumes).

Este vasto conjunto de referência recentes centradas nos instrumentos mostra bem que existe actualmente não apenas uma maior possibilidade de escolha, mas também de opções melhor fundamentadas.

Conclusão

Os instrumentos científicos como o compasso, o astrolábio, a bússola, o microscópio ou o teste psicológico, a título de exemplo, desempenham um papel decisivo em ciência.

Não será difícil ver neste conjunto diversificado de aspectos aqui analisados (os aperfeiçoamentos de natureza técnica na adaptação de instrumentos, a preocupação com as consequências éticas do usos de instrumentos em avaliação psicológica, a extensão da biblioteca básica) vários sintomas.

Em primeiro lugar, constatamos na caracterização que temos vindo a fazer que os instrumentos constituem um domínio com alguma autonomia dentro da avaliação psicológica e incorporam a própria definição estratégica da teoria e prática da avaliação e investigação psicológicas. Nesta linha, os instrumentos introduzem racionalidade e coerência ao processo de aquisição de dados, têm um papel constitutivo regulador na aquisição de informações e na produção de conhecimentos, viabilizam a avaliação psicológica como uma actividade científica sujeita a um trabalho de validação e a uma exigência de prova. Dito de outro modo, os instrumentos devem estar não apenas ao serviço da prática da avaliação, mas também da investigação psicológica.

Depois, observamos a disponibilização e actualização de um cada vez maior número de instrumentos de avaliação, convenientemente estudados e usados em diferentes contextos. Neste plano, reconhece-se que cada instrumento tem uma identidade e desenvolvimento próprios que é, ao mesmo tempo, inseparável do contexto das suas utilizações. O aperfeiçoamento de cada instrumento constitui uma exigência simultaneamente técnica e ética, apoia-se quer nos dados de natureza empírica, quer nos desenvolvimentos teóricos e pode contribuir decisivamente para o progresso da investigação nas várias áreas de exercício da avaliação psicológica.

Finalmente, uma concepção desmistificadora da avaliação psicológica e dos resultados num qualquer instrumento de medida obriga a considerar as pontuações como aproximações de natureza probabilística. Dito de outro modo, o manual da prova deve precisar de forma quantificada os níveis de conhecimento aproximado associados a cada pontuação através da identificação do erro da medida e dos respectivos intervalos de confiança dos resultados. O reconheci-

mento explícito e formal do erro de medida associado a cada pontuação, a recusa de um valor quantitativo único absoluto para caracterizar o desempenho ou funcionamento, constituem uma estratégia técnica que assegura uma maior objectividade e sentido ético ao uso de instrumentos em avaliação psicológica.

A ideia de uma descrição ou explicação completa do funcionamento psicológico do sujeito formulada através dos resultados quantitativos nos instrumentos de avaliação releva de um empirismo ou realismo ingênuos, das ilusões do “absoluto”. Como escreve Bachelard (1996/1938) não há factos “brutos”: “nada é dado, tudo é construído”. Por isso, a “verdade do sujeito” nunca é fácil de conhecer, sobretudo quando a concepção de verdade pensada remete menos para a coerência lógica da caracterização e mais para a adequação ou correspondência exacta entre o discurso construído pelo psicólogo acerca do sujeito e a realidade do seu funcionamento psicológico.

O recurso necessário a instrumentos não é um problema puramente técnico. A prudência, tolerância, humildade, paciência, rigor e disciplina são valores ou princípios da ciência. Por isso, a verdade científica obtida a partir dos instrumentos não é um dado definitivamente adquirido e, nessa medida, deve ser questionada e aperfeiçoada. Neste plano é essencial o domínio dos critérios de cientificidade, a compreensão da lógica interna subjacente ao uso de instrumentos de avaliação psicológica e o conhecimento da força e limites explicativos dos seus resultados.

Referências

- Ackerman, M. J. (1999). *Essentials of forensic psychological assessment*. New York: Wiley.
- Actas dos Congressos “Avaliação Psicológica: Formas e Contextos” [1993-1999, 2004, 7 volumes]. Braga: APPORT/Instituto de Educação e Psicologia da Universidade do Minho.
- Almeida, L. S., Simões, M. R. e Gonçalves, M. (Eds.) (1995). *Provas psicológicas em Portugal*. Braga: APPORT.
- Almeida, L., Simões, M. R., Machado, C. e Gonçalves, M., (Eds.) (2004). *Avaliação Psicológica: Instrumentos validados para a população portuguesa* (Vol. II). Coimbra: Quarteto.
- American, Psychological Association (1999). *Standards for educational and psychological tests*. Washington, DC: APA.
- American, Psychological Association (2002). Ethical principles of psychologists and code of conduct. *American Psychologist*, 47, 1597-1611. [<http://www.apa.org/ethics/stmn01.htm>]

- Andrews, J. L., Saklofski, D. e Janzen, H. (Eds.) (2001). *Handbook of psychoeducational assessment*. New York: Academic Press.
- Archer, R. P., Maruish, M., Imhof, E. A., e Piotrowski, C. (1991). Psychological test usage with adolescents: 1990 survey findings. *Professional Psychology: Research and Practice*, 22, 247-252.
- Bachelard, G. (1996/1938). *La formation de l'esprit scientifique: Contribution à une psychanalyse de la connaissance objective*. Paris: Vrin.
- Barak, A. e English, N. (2002). Prospects and limitations of psychological testing on the Internet. *Journal of Technology in Human Services*, 19, 65-89.
- Barron, J. W. (Ed.) (1998). *Making diagnosis meaningful: Enhancing evaluation and treatment of psychological disorders*. Washington, DC: American Psychological Association.
- Barthélémy, J.-H. (2004). Des instruments de connaissance. *Sciences et Avenir (Hors-Série: Ces instruments qui font la science)*, 140, 46-50.
- Bartram, D. (2001). Guidelines for test users: A review of national and international initiatives. *European Journal of Psychological Assessment*, 17(3), 173-186.
- Bartram, D. e Coyne, I. (1998). Variations in national patterns of testing and test use: The ITC/EFPA international survey. *European Journal of Psychological Assessment*, 14, 249-260.
- Beaune, J.-Cl. (2004). Le génie de la technique: Instrument scientifique et instrument technique. *Sciences et Avenir (Hors-Série: Ces instruments qui font la science)*, 140, 4-7.
- Bénony, H. (2002). *L' examen psychologique et clinique de l'adolescent*. Paris: Nathan.
- Beutler, L. E. e Malik, M. L. (2002). *Rethinking the DSM: A psychological perspective*. Washington, DC: American Psychological Association.
- Beutler, L. E. e Rosner, R. (1995). Introduction to psychological assessment. In L. E. Beutler e M. R. Berren (Eds.), *Integrative assessment of adult personality* (pp. 1-93). New York: Guilford.
- Bouvard, M. (1999). *Questionnaires et échelles d'évaluation de la personnalité*. Paris: Masson.
- Bouvard, M. e Cottraux, J. (1998). *Protocoles et échelles d'évaluation en psychiatrie et en psychologie* (2^a ed.). Paris: Masson.
- Braun, H. I., Jackson, D. N. e Wiley, D. E. (Eds.) (2002). *The role of constructs in psychological and educational measurement*. Mahwah, NJ: Erlbaum.
- Buchanan, T. (2002). Online assessment: Desirable or dangerous? *Professional Psychology: Research and Practice*, 33, 148-154.
- Buela-Casal, G., Caballo, V. E. e Sierra, J. C. (Dirs.) (1996). *Manual de evaluación en psicología clínica y de la salud*. Madrid: Siglo XXI.

- Burns, A., Lawlor, B. e Craig, S. (2004). *Assessment scales in old age psychiatry* (2.^a ed.). London: Martin Dunitz.
- Butcher, J. N., Perry, J. e Hahn, J. (2004). Computers in clinical assessment: Historical developments, present status, and future challenges. *Journal of Clinical Psychology*, 60(3), 331-345.
- Butcher, J. N., Perry, J. N. e Atlis, M. M. (2000). Validity and utility of computer-based test interpretation. *Psychological Assessment*, 12, 6-18.
- Camara, W. J., Nathan, J. S., e Puente, A. E. (2000). Psychological test usage: Implications in professional psychology. *Professional Psychology: Research and Practice*, 31, 141-154.
- Cates, J. A. (1999). The art of assessment in psychology: Ethics, expertise, and validity. *Journal of Clinical Psychology*, 55(5), 631-641.
- Cicchetti, D. V. e Rourke, B. P. (Eds.) (2004). *Methodological and biostatistical foundations of clinical neuropsychology and medical and health disciplines* (2^a ed.). London: Taylor & Francis Group.
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6, 284-290.
- Clark, L. A. e Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7, 309-319.
- Cléro, J.-P. (2004). Le déni de la technique. *Sciences et Avenir (Hors-Série: Ces instruments qui font la science)*, 140, 70-74.
- Cushman, L. A. e Scherer, M. J. (Eds.) (1995). *Psychological assessment in medical rehabilitation*. Washington, DC: American Psychological Association.
- Dana, R. H. (2005). *Multicultural assessment: Principles, applications and examples*. Mahwah, NJ: Erlbaum.
- Davidson, K.W. (1996). Self-and expert-reported emotion inhibition: On the utility of both data sources. *Journal of Research in Personality*, 30, 535-549.
- Debray, R. (2000). *L' examen psychologique de l'enfant à la periode de latence*. Paris: Dunod.
- Dorfman, W. I. & Hersen, M. (Eds.) (2001). *Understanding psychological assessment*. New York: KluwerAcademic/Plenum.
- Downing, S. e Haladyna, T. M. (Eds.) (no prelo). *Handbook of test development*. Mahwah, NJ: Erlbaum.
- Dragow, F. e Olson-Buchanan, J. B. (Eds.) (1999). *Innovations in computerized assessment*. Mahwah, NJ: Erlbaum.
- Edenborough, R. (2005). *Assessment methods in recruitment, selection and performance: A manager's guide to psychometric testing, interviews and assessment centres*. New York: Kogan Page.

- Eignor, D. (2001). Standards for the development and use of tests: The standards for educational and psychological testing. *European Journal of Psychological Assessment*, 17(3).
- Eisman, E. J., Dies, R. R., Finn, S. E., Eyde, L. D., Kay, G. G., Kubiszyn, T. W., Meyer, G. J. e Moreland, K. (2000). Problems and limitations in the use of psychological assessment in the contemporary health care delivery system. *Professional Psychology: Research and Practice*, 31, 131-140.
- Embretson, S.E. e Reise, S. (2000). *Psychometric methods: Item response theory for psychologists*. Mahwah, NJ: Erlbaum.
- Emmanuelli, M. (2004). *L'examen psychologique en clinique*. Paris: Dunod.
- Eyde, L. D., Robertson, G. J., Krug, S. E., Moreland, K. L., Robertson, A. G., Shewan, C. M., Harrison, P. L., Porch, B. E., Hammer, A. L. e Primoff, E. S. (1993). *Responsible test use: Case studies for assessing human behaviour*. Washington, DC: American Psychological Association.
- Fernández-Ballesteros, R. (2003). *Encyclopedia of Psychological Assessment* (Vols. I & II). London: Sage.
- Fernández-Ballesteros, R., De Bruyn, E.E.J., Godoy, A., Hornke, L.F., Ter Laak, J., Vizcarro, C., Westhoff, K., Westmeyer, H. e Zaccagnini, J.L. (2001). Guidelines for the Assessment Process (GAP): A proposal for discussion. *European Journal of Psychological Assessment*, 17(3), 187-200.
- Fisher, C. T. (1992). Humanizing psychological assessment. *The Humanistic Psychologist*, 20, 318-331.
- Foster, S.L. e Cone, J.D. (1995). Validity issues in clinical assessment. *Psychological Assessment*, 7, 248-260.
- Franzen, M. D. (2000). *Reliability and validity in neuropsychological assessment* (2ª ed.). New York: Kluwer Academic/Plenum Publishers.
- Fredericksen, N., Mislevy, R.J. e Bejar, I.I. (Eds.) (1993). *Test theory for a new generation of tests*. Hillsdale, NJ: Erlbaum.
- Garb, H. N. (1998). *Studying the clinician: Judgment research and psychological assessment*. Washington, DC: American Psychological Association.
- Geisinger, K.F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment*, 6, 304-312.
- Golden, C. J., Espe-Pfeifer, P. e Wachsler-Felder (2000). *Neuropsychological interpretations of objective psychological tests*. New York: Kluwer/Plenum.
- Gómez, F. J. (Coord.) (2001/2003). *Evaluación psicológica forense (4 Vols.)*. Salamanca: Amarú Ediciones.
- Gonçalves, M., Simões, M. R., Almeida, L. e Machado, C. (Eds.) (2003). *Avaliação Psicológica: Instrumentos validados para a população portuguesa* (Vol. I). Coimbra: Quarteto.

- Gopaul-McNicol, S.-A. e Armour-Thomas, E. (Eds.) (2002). *Assessment and culture: Psychological tests with minority populations*. New York: Academic Press.
- Graham, J. R. e Naglieri, J. A. (Eds.) (2003). *Handbook of Psychology (Vol. 10: Assessment psychology)*. New York: Wiley.
- Grisso, T. (2003). *Evaluating competencies: Forensic assessments and instruments* (2ª ed.). New York: Kluwer Academic/Plenum Publishers.
- Groth-Marnat, G. (2000). Visions of clinical assessment: Then, now, and a brief history of the future. *Journal of Clinical Psychology*, 56(3), 349-365.
- Groth-Marnath, G. (Ed.) (2002). *Neuropsychological assessment in clinical practice: A guide to test interpretation and integration*. New York: Wiley.
- Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E. e Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment*, 12, 19-30.
- Gutiérrez, T., Raich, R. M., Sánchez, D. e Deus, J. (2003). *Instrumentos de evaluación en psicología de la salud*. Madrid: Alianza Editorial.
- Hambleton, R. J. (1994). Guidelines for adapting educational and psychological tests. *European Journal of Psychological Assessment*, 10, 229-244.
- Hambleton, R.J. (2001). Guidelines for test translation/adaptation. *European Journal Psychological Assessment*, 17(3), 164-172.
- Hambleton, R. K., Merenda, P. F. e Spielberger, C. D. (Eds.) (2005). *Adapting educational and psychological tests for cross-cultural assessment*. Mahwah, NJ: Erlbaum.
- Heilbrun, K. (2001). *Principles of forensic mental health assessment*. New York: Kluwer Academic/Plenum.
- Heilbrun, K. (2003). *Forensic mental health assessment: A casebook*. New York: Kluwer Academic/Plenum.
- Hersen, M. e Porzelius, L. K. (Eds.) (2002). *Diagnosis, conceptualization, and treatment planning for adults: A step-by-step guide*. Mahwah, NJ: Erlbaum.
- Hersen, M. (Ed.) (2004). *Psychological assessment in clinical practice: A pragmatic guide*. New York: Brunner-Routledge.
- Hodges, J. R. (1995). *Cognitive assessment for clinicians*. Oxford: Oxford University Press.
- Hollin, C. R. (Ed.) (2000). *Handbook of offender assessment and treatment*. New York: Wiley.
- Hunt, W. A. (1946). The future of diagnostic testing in clinical psychology. *Journal of Clinical Psychology*, 2, 311-317.
- International Test Commission (ITC) (1999/2003). *Adaptação portuguesa das Directrizes Internacionais para a utilização de testes*. Lisboa: Cegoc. [Comissão para a Adaptação Portuguesa: Abel Pires, António Menezes Rocha, Leandro Almeida, Maria João Afonso, Maria João Seabra Santos, Mário R. Simões e Rui Bártole Ribeiro].

- Irvine, S. e Kyllonen, P. (Eds.) (2001). *Item generation for test development*. Hillsdale, NJ: Erlbaum.
- Lees-Haley, P. R., Smith, H. H., Williams, C. W. e Dunn, J. T. (1996). Forensic neuropsychological test usage: An empirical study. *Archives of Clinical Neuropsychology*, 11, 45-51.
- Krishnamurthy, R., VandeCreek, L., Kaslow, N.J., Tazeau, Y.N., Miville, M. L., Kernes, R., Stegman, R., Suzuki, L. e Benton, S.A. (2004). Achieving competency in psychological assessment: Directions for education and training. *Journal of Clinical Psychology*, 60(7), 725-739.
- Kubiszyn, T. W., Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R. & Eisman, E. J. (2000). Empirical support for psychological assessment in clinical health care settings. *Professional Psychology: Research and Practice*, 31, 119-130.
- Leitão, L. M. (2004) (Coord.), *Avaliação psicológica em orientação escolar e profissional*. Coimbra: Quarteto.
- Lezak, M. D., Howieson, D. B., Loring, D. W., Hannay, H. J., e Fischer, J. S. (2004). *Neuropsychological assessment* (4.^a ed.). New York: Oxford University Press.
- López, M. M., Álvaro, A. R., Santos, E.P., Sánchez, A. B. e Colomina, A. V. (2002). *Instrumentos de evaluación en salud mental*. Madrid: Pirâmide.
- Lopez, S. J. e Snyder, C. R. (Eds.) (2003). *Positive psychological assessment: A handbook of models and measures*. Washington, DC: American Psychological Association.
- Lubin, B. Larsen, R. M. e Matarazzo, J. D. (1984). Patterns of psychological test usage in the United States: 1935-1982. *American Psychologist*, 39, 451-454.
- Maarse, F. J., Mulder, L. J. M., Brand, A. N. e Akkerman, A. E. (2003). *Clinical assessment, computerized methods, and instrumentation*. Lisse: Sweets & Zeitlinger.
- Marlowe, D. B., Wetzler, S. e Gibbings, E.N. (1992). Graduate training in psychological assessment: What Psy.D.'s and Ph.D.'s must know. *The Journal of Training and Practice in Professional Psychology*, 6(2), 9-18.
- Maruish, M. E. (Ed.) (2004a). *The use of psychological testing for the treatment planning and outcomes assessment (Vol. 2: Instruments for children and adolescents)*. Mahwah, NJ: Erlbaum.
- Maruish, M. E. (Ed.) (2004b). *The use of psychological testing for the treatment planning and outcomes assessment (Vol. 3: Instruments for adults)*. Mahwah, NJ: Erlbaum.
- Matarazzo, J. D. (1990). Psychological assessment versus psychological testing: Validation from Binet to the school, clinic, and courtroom. *American Psychologist*, 45, 999-1017.
- McCaffrey, R. J., Duff, K. e Westervelt, H. J. (Eds.) (2000). *Practitioner's guide to evaluating change with neuropsychological assessment instruments*. New York: Kluwer Academic/Plenum Publishers.

- Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of evidence*. Minneapolis: University of Minnesota Press.
- Meyer, G.J., Finn, S.E., Eyde, L.D., Kay, G.G., Kubiszyn, T.W., Moreland, K.L., Eisman, E.J., e Dies, R.R. (1998). *Benefits and costs of psychological assessment in health-care delivery: Report of the Board of Professional Affairs Psychological Assessment Work Group, Part I*. Washington, DC: American Psychological Association.
- Meyer, G. J., Finn, S. E., Eyde, L., Kay, G. D., Moreland, K. L., Dies, R. R., Eisman, E. J., Kubiszyn, T. W. e Reed, G. M. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist*, 56(2), 128-165.
- Mitrushina, M. N., Boone, K. B., Razani, J., e D'Elia, L. F. (2005). *Handbook of normative data for neuropsychological assessment* (2ª ed.). New York: Oxford University Press.
- Moreland, K. L., Eyde, L. D., Robertson, G. J., Primoff, E. S. e Most, R. B. (1995). Assessment of test user qualifications: A research-based measurement procedure. *American Psychologist*, 50, 14-23.
- Muñiz, J., Bartram, D., Evers, A., Boben, D., Matesic, K., Glabeke, K., Fernández-Hermida, J. R. e Zaal, J.N. (2001). Testing practices in European countries. *European Journal of Psychological Assessment*, 17(3), 201-211.
- Muñiz, J., Prieto, G., Almeida, L. e Bartram, D. (1999). Test use in Spain, Portugal and Latin American Countries. *Europeau Journal of Psychological Assessment* 15, 151-157.
- Murphy, L. L., Plake, B. S., Impara, J. C. e Spies, R. A. (2002). *Tests in Print VI*. Lincoln, NE: University of Nebraska Press.
- Naglieri, J. A., Drasgow, F., Schmidt, M., Handler, L., Prifitera, A., Margolis, A. e Velasquez, R. (2004). Psychological testing on the Internet: New problems, old issues. *American Psychologist*, 59 (3), 150-162.
- Newark, C. S. (1996). *Major psychological assessment instruments* (2ª ed.). Needham Heights, MA: Allyn & Bacon.
- Olson, K. R. (2001). Computerized psychological test usage in APA-accredited training programs. *Journal of Clinical Psychology*, 57(6), 727-736.
- Pinard, G.-F. e Pagani, L. (Ed.) (2001). *Clinical assessment of dangerousness: Empirical contributions*. Cambridge: Cambridge University Press.
- Piotrowski, C. (1999). Assessment practices in the era of managed care: Current status and future directions. *Journal of Clinical Psychology*, 55, 787-796.
- Piotrowski, C. e Belter, R. W. (1999). Internship training in psychological assessment: Has managed care had an impact? *Assessment*, 6, 381-385.
- Piotrowski, C. e Lubin, B. (1990). Assessment practices of health psychologists: Survey of Division 38 clinicians. *Professional Psychology: Research and Practice*, 21, 99-106.
- Piotrowski, C. e Zalewski, C. (1993). Training in psychodiagnostic testing in APA - approved PsyD and PhD clinical psychology programs. *Journal of Personality Assessment*, 61, 394-405.

- Plake, B. S. e Impara, J. C. (2001). *The fourteenth mental measurements Yearbook*. Lincoln, NE: University of Nebraska Press.
- Plake, B. S., Impara, J. C. e Spies, R. A. (2003). *The fifteenth mental measurements yearbook*. Lincoln, NE: University of Nebraska Press.
- Rabin, L. A., Barr, W. B. e Burton, L. A. (2005). Assessment practices of clinical neuropsychologists in the United States and Canada: A survey of INS, NAN, and APA Division 40 members. *Archives of Clinical Neuropsychology*, 20, 33-65.
- Reynolds, C. R. (Ed.) (1998). *Comprehensive Clinical Psychology. Vol. 4: Assessment*. Oxford: Elsevier.
- Ribeiro, J. L. P. (1999). *Investigação e avaliação em psicologia e saúde*. Lisboa: Climepsi.
- Rogers, R. e Shuman, D. W. (2000). *Conducting insanity evaluations* (2ª ed.). New York: Guilford.
- Rozensky, R. H., Sweet, J. J. e Tovian, S. M. (2004). *Psychological assessment in medical settings*. New York: Kluwer.
- Shapiro, E. S. e Kratochwill, T. R. (Eds.) (2000). *Conducting school-based assessments of child and adolescent behavior*. New York: Guilford.
- Shulman, K. e Feinstein, A. (2003). *Quick cognitive screening for clinicians: Mini Mental, Clock Drawing and other brief tests*. New York: Martin Dunitz.
- Simões, M. R. (1994). Notas em torno da arquitectura da Avaliação Psicológica. *Psicologica*, 11, 7-44.
- Simões, M. R., Almeida, L. S. e Gonçalves, M. M. (1999). Testes e Provas Psicológicas em Portugal: Roteiro de algumas questões que atravessam a utilização de instrumentos de/na Avaliação Psicológica. In Mário R. Simões, L. S. Almeida e M. M. Gonçalves (Eds.), *Testes e Provas Psicológicas em Portugal* (vol. II, pp. 1-12). Braga: Associação dos Psicólogos Portugueses/Sistemas Humanos e Organizacionais.
- Simões, M. R., Gonçalves, M. e Almeida, L. (1999). *Testes e Provas Psicológicas em Portugal* (Vol. II). Braga: SHO/APPORT.
- Simões, M. R., Machado, C., Gonçalves, M. & Almeida, L. S. (Coords.) (no prelo). *Avaliação Psicológica: Instrumentos validados para a população portuguesa* (vol. III). Coimbra: Quarteto.
- Simões, M. R. e Lopes, A. F. (2004). Contributos da avaliação para a planificação e medida da intervenção psicológica. In H. Damião da Silva, A. Castro Fonseca, L. Alcoforado, M. Manuela Vilar e Cristina Vieira (Eds.), *Crianças e jovens em risco: Da investigação à intervenção* (pp.165-209). Coimbra: Almedina/Centro de Psicopedagogia.
- Smith, G. T. e McCarthy, D. M. (1995). Methodological considerations in the refinement of clinical assessment instruments. *Psychological Assessment*, 7 (3), 300-308.
- Spies, R. A. e Plake, B. S. (2005). *The sixteenth mental measurements yearbook*. Lincoln, NE: University of Nebraska Press.
- Spren, O. & Strauss, E. (1998). *A compendium of neuropsychological tests* (2.ª ed.). New York: Oxford University Press.

- Stedman, J.M., Hatch, J.P. e Schoenfeld, L.S. (2000). Pre-internship preparation in psychological testing and psychotherapy: What internships directors say they expect. *Professional Psychology: Research and Practice*, 31, 321-326.
- Stout, C. E. (1992). Psychological assessment training in professional schools: Literature review and personal impressions. *The Journal of Training and Practice in Professional Psychology*, 6(1), 14-21.
- Sundberg, N. D. (1961). The practice of psychological testing in clinical services in the United States. *American Psychologist*, 16, 79-83.
- Suzuki, L. A., Ponterotto, J. G. e Meller, P. J. (2001). *Handbook of multicultural assessment: Clinical, psychological and educational applications*. New York: Wiley.
- Sweet, J. J., Moberg, P. J. e Westergaard, C. K. (1996). Five-year follow-up survey of practices and beliefs of clinical neuropsychologists. *The Clinical Neuropsychologist*, 10, 202-221.
- Sweet, J. J., Peck, E. A., Abramowitz, C. e Etzweiler, S. (2002). National Academy of Neuropsychology/Division 40 of the American Psychological Association practice survey of clinical neuropsychology in the United States, Part I: Practitioner and practice characteristics, professional activities and time requirements. *The Clinical Neuropsychologist*, 16, 109-127.
- Thissen, D. e Wainer, H. (Eds.) (2001). *Test scoring*. Mahwah, NJ: Erlbaum.
- Thomas, J. C. (Ed.) (2003). *Comprehensive handbook of psychological assessment (Vol. 4: Industrial and organizational assessment)*. New York: Wiley.
- Turner, S.M., DeMers, S.T., Fox, H.R. e Reed, G.M. (2001). APA's guidelines for test user qualifications. *American Psychologist*, 56, 1099-1113.
- Urbina, S. (2004). *Essentials of psychological testing*. New York: Wiley.
- Wade, D.T. (1995). *Measurement in neurological rehabilitation*. Oxford: Oxford University Press.
- Wainer, H. (2000). *Computerized adaptive testing: A primer* (2.^a ed.). Mahwah, NJ: Erlbaum.
- Waller, N.G., Yonce, L.J., Grove, W.M., Faust, D. e Lenzenweger, M.F. (Eds.) (2005). *A Paul Meehl reader: Essays on the practice of scientific psychology*. Mahwah, NJ: Erlbaum.
- Watkins, C.E., Jr. (1991). What have surveys taught us about the teaching and practice of psychological assessment? *Journal of Personality Assessment*, 56, 426-437.
- Watkins, C.E., Jr., Campbell, V.L., Nieberding, R. e Hallmark, R. (1995). Contemporary practice of psychological assessment by clinical psychologists. *Professional Psychology: Research and Practice*, 26, 54-60.
- Wilson, M. (2005). *Constructing measures*. Mahwah, NJ: Erlbaum.

PSYCHOLOGICAL ASSESSMENT INSTRUMENTS: POTENTIALITIES AND LIMITATIONS

Mário R. Simões

Serviço de Avaliação Psicológica

Faculdade de Psicologia e de Ciências da Educação,

Universidade de Coimbra, Portugal

Abstract: This article reviews the purposes, functions and applications of psychological measurement instruments in the context of psychological assessment practices and research. Some comments are made featuring some of the most salient and controversial topics about the uses of psychological measurement instruments: the technical developments, ethical and interpretation issues and their contributions to the psychological assessment.

KEY-WORDS: *Psychological Assessment, Psychological measurement instruments, Validity, Ethics, Interpretation, Research.*

VISUALIZACIÓN ESPACIAL Y RENDIMIENTO ACADÉMICO EN DIBUJO TÉCNICO

Gerardo Prieto, Rosario Arias-Barahona
Universidad de Salamanca (España)

Ana-María Núñez
Instituto Peruano de Psicología y Educación (Perú)

Resumen

Las dificultades de aprendizaje del Dibujo Técnico que experimentan los estudiantes de Ingeniería se relacionan con su nivel de Aptitud Espacial. Para mejorar el proceso didáctico, sería necesario detectar tempranamente a los estudiantes que requieren más apoyo. Este estudio describe el análisis de un test de Aptitud Espacial que podría facilitar un diagnóstico temprano relacionado con el rendimiento de los estudiantes en Dibujo Técnico. Se aplicó un test de Visualización (TVZ2002-C) a una muestra de estudiantes de Ingeniería de la Universidad de San Martín de Porres (Lima, Perú). Los resultados sugieren que el test puede ser útil para detectar a los estudiantes con diferentes niveles de rendimiento académico.

PALAVRAS-CHAVE: *Aptitud espacial, predicción del rendimiento académico, dibujo técnico, modelo de Rasch.*

La Aptitud Espacial ha sido definida como la capacidad para formar, retener, recuperar y transformar imágenes visuales (Carroll, 1993). En los estudios factoriales se han identificado varias aptitudes espaciales, que se caracterizan diferencialmente por el énfasis en los diferentes aspectos de los procesos implicados: formación, almacenamiento, recuerdo y transformación de la imagen (Lohman, 1994). La aptitud de Visualización es, junto a las de Relaciones Espaciales y Orientación Espacial, uno de los factores más representativos. Ha sido definido como la aptitud para manipular mentalmente figuras tridimensionales complejas. En la ejecución de las tareas empleadas para su diagnóstico, entre las que destaca la denominada "desarrollo de superficies", se pone más énfasis en la precisión que en la rapidez de la respuesta (Juan-Espinosa, 1997).

Morada (address): Gerardo Prieto. Departamento de Psicología Básica, Psicobiología y Metodología. Facultad de Psicología. Universidad de Salamanca. Avenida de la Merced, 109-131. 37005 Salamanca. Tfn.: 923294500-3269. Correo electrónico: gprieto@usal.es.

A pesar de la importancia de la Aptitud Espacial en las teorías sobre la inteligencia, los tests espaciales no han sido muy usados para la selección de personal y el diagnóstico educativo. La excepción a esta regla general aparece fundamentalmente en el contexto de la selección de pilotos y controladores aéreos. La razón principal de este escaso empleo en el ámbito pedagógico radica en que las aptitudes verbales (lectura, escritura, comprensión y expresión oral) están más involucradas en el aprendizaje académico convencional (Kyllonen y Glück, 2003). Sin embargo, existen estudios que manifiestan que las aptitudes espaciales son buenos predictores del rendimiento en materias específicas de las especialidades de arquitectura, ingeniería, diseño gráfico y mecánico (Peters, Chisholm y Laeng, 1995; Hsi, Linn y Bell, 1997; Ting, 2001; Prieto y Velasco, 2002b).

El papel mediador de la Aptitud Espacial en el rendimiento en materias técnicas y artísticas, tales como el Dibujo Técnico, puede explicarse porque los procesos de transformación mental de imágenes visuales han de ser empleados con precisión por los aprendices. Por ejemplo, los estudiantes de estas disciplinas necesitan imaginar la apariencia de los objetos en distintas orientaciones, trasladar a un espacio de tres dimensiones representaciones bidimensionales, imaginar cómo se verían las partes ocultas de un objeto, etc. De hecho, los profesores de Dibujo Técnico manifiestan frecuentemente que algunos alumnos no progresan adecuadamente en el proceso de aprendizaje debido a su bajo nivel de Aptitud Espacial. Es decir, para la solución de problemas no es suficiente emplear estrategias analítico-verbales, también son necesarias las estrategias holístico-espaciales: es necesario “visualizar” mentalmente el problema (Lohman, 2000).

Desde esta perspectiva, sería muy útil el diagnóstico precoz de la Visualización para identificar a los alumnos menos capaces, con el fin de programar métodos de enseñanza específicos para ellos.

En este trabajo hemos analizado la eficacia predictora del rendimiento académico en Dibujo Técnico de un test de Visualización. Nuestro interés, ha estado dirigido a los estudiantes de ingeniería cuya formación le exige el uso de esta aptitud a lo largo de toda la carrera.

Metodo

Participantes

Participaron en el estudio los integrantes de una muestra incidental de 178 estudiantes de la Universidad de San Martín de Porres (Lima, Perú). Los alumnos

estaban matriculados en Ingeniería de Sistemas, Electrónica, Industrial y Civil, y cursaban la asignatura de Dibujo Técnico y Sistema de Graficación. La media de la edad de los alumnos era de 19 años y 10 meses, y la desviación típica de 2 años y 8 meses. El 72,5% de los participantes eran varones.

Medidas

1. Test de Visualización Espacial: TVZ2002-C. Se trata de una prueba extraída de un banco de ítems construido ex profeso para el diagnóstico específico de la Visualización en estudiantes de Ingeniería y Arquitectura (Prieto y Velasco, 2002a y 2004). Los ítems del banco han sido construidos a partir de indicadores cognitivos de la Visualización (Embretson, 1996) y están calibrados con el modelo de Rasch (1960).

La prueba está tipificada entre los tests de lápiz y papel y el material está constituido por un cuadernillo impreso que contiene 20 ítems gráficos, los cuales deben ser respondidos en el tiempo límite de 30 minutos.

Los ítems se basan en una tarea de "Desarrollo de Superficies". La tarea consiste en un cubo dibujado con todas sus caras identificadas por letras. A su derecha, el cubo se muestra desdoblado con una de sus caras identificada y otra marcada con signo de interrogación (?). El examinado ha de identificar qué letra, en su posición relativa, corresponde en la cara donde se ubica el signo de interrogación. El examinado tiene que escoger la respuesta correcta entre nueve alternativas (Figura 1).

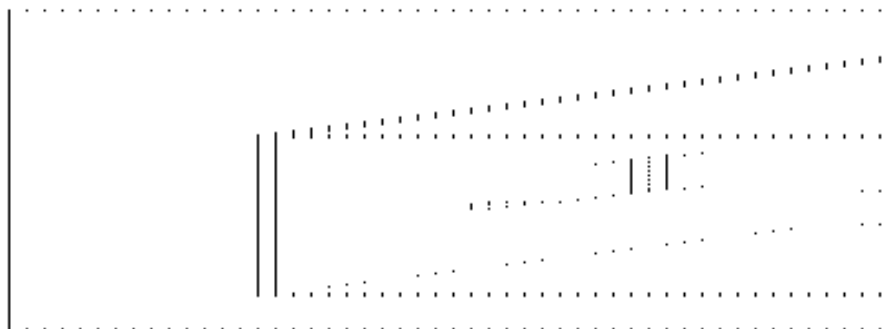


Figura 1. Modelo de ítem del TVZ-2002. La respuesta correcta está marcada con la flecha

Estudios con versiones similares de este test muestran que es un indicador válido y confiable del constructo de Visualización (Prieto y Velasco, 2002b y 2004).

2. Examen académico oficial de la asignatura

El curso de Dibujo Técnico y Sistemas de Graficación es un curso teórico-práctico orientado a la representación gráfica de objetos. El propósito del cur-

so es brindar al estudiante los conocimientos básicos de representación de objetos, que aunados a sus conocimientos previos de matemáticas, geometría plana y geometría analítica, le aporten una herramienta moderna que le permita, con la utilización de una computadora y un software CAD, la elaboración de planos de excelente calidad. Contenidos del curso: Editor de dibujo de AutoCAD 2000 en el entorno Windows; Sistemas de coordenadas; Construcciones geométricas; Geometría aplicada; Capas; Acotado; Dibujos isométricos; Dibujo de objetos en 3D; Visualización de sólidos y Proyecciones múltiples o diédricas. La calificación en el examen final, basado en solución de problemas relativos a los contenidos mencionados, se sitúa en una escala de 0 a 20 puntos. La nota mínima para superar el curso es 11.

Procedimiento

Se administró el TVZ-2002-C, al inicio del semestre académico (Agosto 2004), a los integrantes de la muestra. La administración fue colectiva en grupos no mayores de 25 alumnos. El examen final tuvo lugar al final del semestre (Noviembre 2004).

Modelo de Medida

Se analizaron los datos mediante un modelo TRI muy conocido: el modelo logístico de un parámetro (Rasch, 1960). La ecuación básica (1) especifica que la probabilidad de resolver correctamente un ítem depende solamente de la diferencia en el atributo medido entre el nivel de la persona (θ_s) y el nivel del ítem (β_j). En este caso,

$$P_{js} = \frac{e^{(\theta_s - \beta_j)}}{1 + e^{(\theta_s - \beta_j)}} \quad (1)$$

Donde e es la base de los logaritmos naturales (2,7183).

Los valores escalares de las personas y los ítems se sitúan en la misma escala. Estos valores pueden expresarse en distintas métricas (Embretson y Reise, 2000). La más utilizada es la escala *logit*, que es el logaritmo natural de $(P_{js} / 1 - P_{js})$. Aunque la escala *logit* puede adoptar valores entre más y menos infinito, la gran mayoría de los casos se sitúa en el rango ± 5 . La localización del punto 0 de la escala es arbitraria. En la tradición de Rasch, se suele situar dicho punto en la dificultad media de los ítems.

El modelo de Rasch tiene unas propiedades métricas óptimas: estadísticos suficientes, objetividad específica, unidimensionalidad, métrica intervalar, parámetros independientes de la muestra y estadísticos para analizar el ajuste al modelo de ítems y personas (Santisteban y Alvarado, 2001). Aunque la fun-

cionalidad de estas ventajas depende del ajuste de los datos a los supuestos del modelo, es conveniente notar que basta con un ajuste razonable para justificar su uso: los estudios acerca de la robustez del modelo de Rasch ante las violaciones de sus supuestos muestran que las ganancias producidas por el uso de otros modelos más complejos son escasas (Muñiz, 1997).

Para analizar el ajuste de los datos al modelo y estimar los parámetros de los ítems y las personas, se utilizó el programa Winsteps (Wright y Linacre, 1998).

Como indicador del ajuste de ítems y personas se han empleado los estadísticos Infit y Outfit, que son medias de los residuos estandarizados al cuadrado (diferencias entre el valor observado y el predicho por el modelo). La fórmula de un residuo es:

$$y_{is} - x_{is} P_{is} \quad (2)$$

Donde x_{is} es la respuesta observada y P_{is} la probabilidad de una respuesta correcta de la persona s al ítem i . Se suelen estandarizar los residuos dividiéndolos por su desviación típica:

$$Z_{is} = (x_{is} - P_{is}) / \sqrt{P_{is}(1 - P_{is})} \quad (3)$$

Infit es la media de los residuos cuadráticos ponderados con su varianza (Wis).

$$\text{Infit} = \sum Z_{is}^2 W_{is} / \sum W_{is} \quad (4)$$

Outfit es la media de los residuos cuadráticos no ponderados.

$$\text{Outfit} = \sum Z_{is}^2 / N \quad (5)$$

El valor esperado de estos estadísticos es 1. Por convención se considera que los valores superiores a 1,5 indican un desajuste moderadamente alto, y que los valores superiores a 2,0 revelan muy alto desajuste (Wright y Linacre, 1998).

Outfit es un indicador muy sensible a los outliers (basta una respuesta muy inesperada para que adopte un valor muy elevado). Infit es más robusto: los valores altos se deben a patrones de respuesta claramente aberrantes.

Como medida de la fiabilidad de las puntuaciones se ha empleado un análogo del coeficiente de fiabilidad clásico: el índice de fiabilidad (6). Este estadístico se calcula por separado para las personas y para los ítems.

Fiabilidad = varianza observada - varianza de error / varianza observada (6)

Resultados

En primer lugar, presentaremos los resultados del análisis del ajuste al modelo de los ítems y de los participantes (Tabla 1). Como ya se ha comentado, el ajuste es crucial; en su ausencia, los valores carecen de significado teórico y las ventajas del modelo de Rasch se desvanecen. Los datos manifiestan un buen ajuste al modelo: por un lado, las medias y las desviaciones típicas de los valores de Infit y Outfit son las que cabe esperar cuando no hay divergencias sustanciales entre las predicciones del modelo y los datos empíricos; por otro, no aparecen ítems y sólo 8 personas con valores de Infit y/o Outfit mayor de 2.

Tabla 1. *Estadísticos de ajuste de los ítems y de las personas*

Estadístico	Infit (Ítems)	Outfit (Ítems)	Infit (Personas)	Outfit (Personas)
Media	0,99	1,05	1,00	1,05
D.T.	0,12	0,24	0,16	0,48
Máximo	1,20	1,60	1,41	3,16
N y (%) > 1,5	0 (0,00)	1 (5,00)	0 (0,00)	21 (11,80)
N y (%) > 2,0	0 (0,00)	0 (0,00)	0 (0,00)	8 (4,49)

En la Tabla 2 aparecen los estadísticos descriptivos de las puntuaciones en la escala de los ítems y las personas. Puede observarse que se ha medido a las personas y a los ítems con alta fiabilidad. Aunque el test es difícil para la muestra (la media de las personas es sensiblemente inferior a la dificultad media de los ítems), registra una alta variabilidad en los alumnos.

Tabla 2. *Estadísticos de las puntuaciones de los ítems y de las personas*

Estadístico	Ítems	Personas
Media	0,00	- 0,85
D.T.	0,71	1,64
Máximo	1,31	4,42
Mínimo	- 1,33	- 4,41
Fiabilidad	0,92	0,84

En consecuencia, se puede considerar que las puntuaciones tienen óptimas propiedades métricas y son adecuadas para evaluar las diferencias en el constructo de Visualización.

Los estudios sobre diferencias de género en la Aptitud Espacial informan de una diferencia consistente a favor de los varones (Halpern, 1992). Esta diferencia también aparece en las puntuaciones del test TVZ2002-C. Los alumnos presentan en promedio puntuaciones mayores ($M = -0,66$; $DT = 1,64$; $N = 129$) que las alumnas ($M = -1,35$; $DT = 1,56$; $N = 49$). La diferencia es significativa ($t = 2,58$; $gl = 176$; $p < 0,01$) y el tamaño del efecto “medio” ($0,44$) según el criterio de Cohen (Rosenthal y Rosnow, 1991).

Con la finalidad de analizar la utilidad del test TVZ2002-C para pronosticar el rendimiento académico en Dibujo Técnico y Sistemas de Graficación, se estimó la correlación producto-momento de Pearson entre ambas variables. Dicha correlación ($0,35$) revela una relación moderada. Esta correlación implica que a partir de las puntuaciones del test se puede explicar un 12,25% de la varianza de las calificaciones.

Asimismo, se cuantificaron las medias en el test de los alumnos que obtuvieron en el examen final de la asignatura un rendimiento insuficiente (calificaciones inferiores a 11) y un rendimiento aceptable (calificaciones iguales o mayores que 11). Se llevó a cabo un contraste t de Student de la diferencia entre las medias (Tabla 3).

Tabla 3. Nota en el examen final (rendimiento) y media en el test TVZ2002-C

Nota	Rendimiento	N	Media	DT	T Student	gl	p	Tamaño del efecto (d de Cohen)
$X \geq 11$	Suficiente	97	-0,46	1,63	2,88	152	0,01	0,48
$X < 11$	Insuficiente	57	-1,23	1,54	--	--	--	--

Puede observarse que los alumnos con fracaso académico en la asignatura tienen una media en las puntuaciones del test, aplicado a principio del curso, que es menor que la de los alumnos que superaron la materia a final de curso. La diferencia entre ambas medias es significativa y la magnitud del efecto es de tipo medio (media desviación típica) según la taxonomía de Cohen (Rosenthal y Rosnow, 1991).

Discusión y Conclusión

La Aptitud Espacial tiene una indudable importancia teórica en el contexto de la psicología científica, pero no ha ocupado la misma jerarquía en el contexto aplicado. Por una parte, los modelos jerárquicos de la inteligencia humana (Carroll, 1993) colocan el factor verbal y el factor espacial en el estrato

que se sitúa inmediatamente después de la inteligencia general, puesto que estos factores capturan más varianza de las baterías de tests aptitudinales que cualquiera de las otras dimensiones. Asimismo, las investigaciones neuropsicológicas sugieren que la diferencia entre el procesamiento analítico-verbal y el analógico-espacial es una dicotomía imprescindible para comprender la cognición humana (Lohman, 1996). Por otra parte, existen numerosos estudios acerca de la predicción del rendimiento académico en los que se han utilizado una gran variedad de predictores, tales como la inteligencia, el rendimiento escolar previo, los intereses, los rasgos de personalidad, la motivación, el clima escolar, la clase social, etc. En estos estudios se ha demostrado que, cuando se controla el efecto de la relación entre la inteligencia y el resto de los predictores, el nivel intelectual es el mejor predictor del rendimiento (Brody, 1992; Jensen, 1980; Juan-Espinosa, 1997).

Sin embargo, pese a la importancia teórica de la Aptitud Espacial y a su probable potencial aplicado en el ámbito educativo, los tests espaciales no han sido suficientemente usados. La explicación más plausible es la mayor efectividad del factor verbal y del factor de razonamiento para pronosticar el rendimiento académico en los programas convencionales. No obstante, algunos estudios publicados en la primera mitad del pasado siglo (Holliday, 1943) ya manifestaban la importancia de la Aptitud Espacial para pronosticar el rendimiento académico en estudios técnicos, especialmente en el ámbito de la ingeniería y la arquitectura. En este trabajo, hemos tratado de retomar este tema poniendo el énfasis en la utilidad de un test de Visualización Espacial para detectar a comienzo del curso a los alumnos con probable fracaso académico en la asignatura de Dibujo Técnico.

Los datos obtenidos revelan que la calidad psicométrica de las puntuaciones del test es adecuada: se ajustan satisfactoriamente al modelo de Rasch, considerado como un modelo de medida "fuerte" en la Psicometría moderna (Embretson y Hershberger, 1999) y son altamente fiables. Además, presentan evidencias sobre su validez, tales como un impacto del género. Es decir, los chicos tienen en promedio un nivel superior que las chicas en el test TVZ2002-C. El tamaño del efecto es "medio" de acuerdo con la clasificación de Cohen. Este dato converge con todos los estudios publicados acerca de las diferencias entre sexos en Aptitud Espacial (Halpern, 1992). Aunque el tamaño del efecto es algo mayor que el que aparece en los meta-análisis acerca de las diferencias en Visualización, se ha de notar que el TVZ2002-C tiene un fuerte componente de rotación mental y es en este proceso en que se manifiestan las mayores diferencias de género.

La correlación entre las puntuaciones de los alumnos en el TVZ2002-C y el examen de final de curso en la asignatura de Dibujo Técnico y Sistemas de Graficación es de 0,35, magnitud semejante a la obtenida en otros estudios (Peters, Chisholm y Laeng, 1995; Hsi, Linn y Bell, 1997; Ting, 2001; Prieto y Velasco, 2002b). Dado que en la población universitaria, especialmente en la carrera de ingeniería, hay una fuerte restricción del rango producido por la heteroselección a lo largo del currículum y la autoselección de los alumnos (Hunter y Schmidt, 1990), esta correlación revela una asociación moderadamente alta. Como afirma Juan-Espinosa (1997), "la correlación entre la inteligencia y el rendimiento académico, que normalmente es de 0,50 durante la etapa escolar normal, puede verse reducida incluso hasta 0,1 en poblaciones universitarias".

Otro argumento favorable a la buena eficacia predictora del TVZ2002-C se basa en su utilidad para discriminar entre los alumnos que superaron o fracasaron en el examen final de la asignatura: los primeros obtuvieron un promedio mayor en el test, administrado a comienzo del curso, que los segundos. La diferencia es significativa y el tamaño del efecto es "medio", de acuerdo con el criterio de Cohen ya citado.

Estos datos revelan que el test podría ser empleado rutinariamente como una herramienta para un diagnóstico precoz de los alumnos con dificultades de aprendizaje, a fin de dedicarles mayor apoyo, atención y métodos didácticos específicos.

Referências

- Brody, N. (1992). *Intelligence*. San Diego, CA: Academic Press.
- Carroll, J.B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
- Embretson, S.E. (1996). Cognitive design principles and the successful performer: A study on spatial ability. *Journal of Educational Measurement*, 33, 29-39.
- Embretson, S. e Hershberger, S.L. (1999). *The new rules of measurement*. Mahwah, NJ: LEA.
- Embretson, S.E. e Reise, S.P. (2000) *Item response theory for psychologists*. Mahwah, NJ: LEA.
- Halpern, D.F. (1992). *Sex differences in cognitive abilities*. Hillsdale, NJ: LEA.
- Holliday, F. (1943). The relations between psychological test scores and subsequent proficiency of apprentices in the engineering industry. *Occupational Psychology*, 17, 168-185.
- Hsi, S., Linn, M.C. e Bell, J.E. (1997). The role of spatial reasoning in engineering and design of spatial instruction. *Journal of Engineering Education*, 84, 151-158.
- Hunter, J.E. e Schmidt, F.L. (1990). *Methods of meta-analysis: Correction error and bias in research findings*. Newbury Park: Sage.
- Jensen, A.R. (1980). *Bias in Mental Testing*. New York: Free Press.
- Juan-Espinosa, M. (1997). *Geografía de la inteligencia humana. Las aptitudes cognitivas*. Madrid: Ediciones Pirámide.
- Kyllonen, P.C. e Glück, J. (2003). Spatial ability: Introduction to the special issue. *International Journal of Testing*, 3, 215-217.
- Lohman, D. F. (1994). Spatial ability. In R. Sternberg (Ed.), *Encyclopedia of human intelligence* (pp. 1000-1007). New York: MacMillan.
- Lohman, D.F. (1996). Spatial ability and g. In I. Dennis e P. Tapsfield (Eds.), *Human Abilities: Their nature and measurement* (pp. 97-116). Mahwah, NJ: LEA.
- Lohman, D.F. (2000). Complex information processing and intelligence. In R.J. Sternberg (Org.), *Handbook of intelligence* (pp. 285-340). Cambridge, UK: Cambridge University Press.
- Muñiz, J. (1997). *Introducción a la teoría de respuesta a los items*. Madrid: Pirámide.
- Peters, M., Chisholm, P. e Laeng, B. (1995). Spatial ability, student gender, and academic performance. *Journal of Engineering Education*, 84, 69-73.
- Prieto, G. e Velasco, A. D. (2002a). Construção de um teste de visualização a partir da psicologia cognitiva. *Avaliação Psicológica*, 1, 39-47.

- Prieto, G. e Velasco, A.D. (2002b). Predicting academic success of engineering students in technical drawing from visualization test scores. *Journal for Geometry and Graphics*, 6, 99-109.
- Prieto, G. e Velasco, A. D. (2004). Training visualization ability y technical drawing. *Journal for Geometry and Graphics*, 8, 107-115.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.
- Rosenthal, R. e Rosnow, R.L. (1991). *Essentials of behavioral research: Methods and data analysis*. London: McGraw-Hill.
- Santisteban, C. e Alvarado, J.M. (2001). *Modelos psicométricos*. Madrid: UNED.
- Ting, S.R. (2001). Predicting academic success of first-year engineering students from standardized test scores and psychosocial variables. *International Journal of Engineering Education*, 17, 75-80.
- Wright, B.D. e Linacre, J.M. (1998). *WINSTEPS: A Rasch computer program*. Chicago: MESA Press.

SPATIAL VISUALIZATION AND ACADEMIC ACHIEVEMENT IN TECHNICAL DRAWING

Gerardo Prieto, Rosario Arias-Barahona
Universidad de Salamanca (España)

Ana-María Núñez
Instituto Peruano de Psicología y Educación (Perú)

Abstract: The difficulties of First-Year Engineering students toward learning Technical Drawing are related with a insufficient level in visuo-spatial ability. In order to effectively plan the didactic process, it is necessary to detect those students who require more attention and support. This study proposes an analysis of a psychometric test that could facilitate an early diagnostic of visuo-spatial ability related with the academic performance of students in Technical Drawing. To this end, a Visualization test (TVZ2002-C) was carried out on a sample of Engineering students from the San Martín de Porres University from Lima (Peru). The test was analysed by the Rasch Model. The results suggest that TVZ2002-C can be useful in detecting those students with different performance levels in Technical Drawing.

KEY-WORDS: *Spatial aptitude, predicting academic achievement, technical drawing, Rasch model.*

APTIDÕES COGNITIVAS E RENDIMENTO ACADÉMICO: A VALIDADE PREDITIVA DOS TESTES DE INTELIGÊNCIA

Leandro S. Almeida e Gina Lemos

Departamento de Psicologia, Instituto de Educação e Psicologia, Universidade do Minho

Resumo

Com base numa amostra de alunos portugueses do 5º e 6º anos de escolaridade (n=1163), analisamos a validade dos resultados na Bateria de Provas de Raciocínio (BPR5/6; Almeida, 2003) por referência ao rendimento escolar dos alunos, considerando quer as suas classificações em várias disciplinas curriculares quer o índice de retenções escolares anteriores. De uma maneira geral os resultados sugerem uma associação com significado estatístico entre rendimento escolar e desempenho nas quatro provas de raciocínio da bateria (raciocínio abstracto, raciocínio verbal, raciocínio numérico e resolução de problemas). Os alunos com melhor rendimento na escola apresentam também melhores resultados na BPR5-6, o que vai no sentido da sua validade por referência a critério externo. Os resultados nas provas de resolução de problemas e de raciocínio verbal apresentam-se mais correlacionados com o rendimento escolar, ao mesmo tempo que esses coeficientes são mais elevados quando tomamos as disciplinas curriculares mais tipicamente académicas (Português, Matemática, Ciências da Natureza). Os resultados mostram, ainda, que medidas globais em relação às disciplinas curriculares e em relação às quatro provas da bateria se apresentam mais correlacionadas entre si, atingindo .65 para os alunos do 5º ano e .63 para os alunos do 6º ano de escolaridade. Finalmente, os resultados nas provas cognitivas encontram-se associados às retenções escolares dos alunos. Os alunos sem qualquer retenção escolar anterior atingem médias mais elevadas nas provas da BPR5-6. Este conjunto de resultados obtidos apoiam a validade externa da BPR5-6 e apoiam a sua utilização pelos psicólogos escolares nas dificuldades de aprendizagem dos alunos.

PALAVRAS-CHAVE: *Inteligência, Rendimento escolar, Validade preditiva, Testes de raciocínio, Avaliação psicológica.*

Introdução

A investigação em Portugal e noutros países tem evidenciado uma correlação positiva e estatisticamente significativa entre as classificações escolares dos alunos e as suas capacidades cognitivas (Almeida, 1988a; Almeida e

Campos, 1986; Jensen, 1998; Neisser, Boodoo, Bouchard, Boykin, Brody, Ceci, Halpern, Loehlin, Perloff, Sternberg, e Urbina, 1996; Primi e Almeida, 2000; Schmidt e Hunter, 1998; Schmidt, Pearlman e Hunter, 1980). Esta informação tem servido o objectivo de validade dos próprios testes de inteligência ou aptidões, acabando por ser também uma informação de elevada importância na compreensão das dificuldades de aprendizagem por parte dos psicólogos escolares.

De entre vários constructos psicológicos, a inteligência permanece como um aspecto importante no contexto educativo (Almeida, 1992, 1996ab; Pereira e Simões, 2005; Te Nijenhuis, Evers, e Mur, 2000; Te Nijenhuis, Tolboom e Bleichrodt, 2004; Simões et al., 1998, 2003). Os testes de inteligência, e em particular os testes de factor g e as escalas de QI, assumem-se como bons preditores do sucesso escolar (Almeida, 1988a, 1996b; Simões e Albuquerque, 2002; Te Nijenhuis et al., 2004). Os coeficientes de correlação entre testes de inteligência e notas escolares aproximam-se de .50 (Jensen, 1998; Neisser et al., 1996; Snow e Yalow, 1982). Para Jensen (1980, p.317) as crianças com mais elevado QI aprendem mais e melhor, aprendem mais depressa e permanecem na escola por mais anos. Este facto, associado à validade dos próprios testes de inteligência, explica a continuidade do seu uso na educação enquanto um dos critérios mais simples e que melhor prediz o rendimento escolar dos alunos (Te Nijenhuis et al., 2004).

Apesar da relação estabelecida entre variáveis cognitivas e resultados da aprendizagem, os coeficientes de correlação obtidos podem oscilar ao longo da escolaridade. Os coeficientes são mais baixos no secundário e superior em virtude das amostras se tornarem mais homogêneas. Também oscilam segundo a natureza das provas cognitivas usadas. Da nossa própria experiência, uma combinação de resultados em várias provas cognitivas, mais ou menos diferenciadas, pode permitir uma boa estimativa do nível de aptidão geral dos alunos e melhor abarcar as suas capacidades e funcionamento cognitivo, acabando por isso por atingirem níveis mais elevados de correlação com as classificações escolares dos alunos (Almeida, 1988a; Almeida e Campos, 1986; Bleichrodt, Resing, Drenth e Zaal, 1984, 1987; Meuris, Almeida e Campos, 1988).

Assumida a relação entre habilidades cognitivas e rendimento escolar, tem sido usual utilizar-se as classificações académicas dos alunos como um dos critérios externos para a validação dos resultados nos testes de inteligência. Este objectivo tem explicado um recurso usual desta estratégia de validação dos testes de inteligência, mesmo que para alguns autores se possa dizer que ele pode reforçar uma menor evolução do método dos testes ou que tal

reforça uma “inteligência escolástica” dominante na generalidade dos testes de inteligência (Almeida, 1994; Sternberg e Kaufman, 1996).

Independentemente desta discussão, a ligação entre resultados nos testes de inteligência e rendimento escolar tem permitido aos psicólogos, nas suas práticas, buscarem no desenvolvimento e no funcionamento cognitivo dos alunos alguma informação relevante para explicar os bons e fracos desempenhos académicos dos estudantes. Acreditando-se no substracto cognitivo comum à aprendizagem e à inteligência (Almeida, 1992, 1996a), tradicionalmente associaram-se as dificuldades na aprendizagem às menores capacidades cognitivas dos alunos, servindo os testes de inteligência um objectivo de diagnóstico. Ao mesmo tempo, quando se lê a inteligência de um modo multifacetado (aptidões), a sua avaliação pode também servir a intervenção psicológica de apoio às escolhas vocacionais dos alunos. Nomeadamente nos momentos de transição escolar envolvendo as escolhas pelos alunos de disciplinas e/ou cursos, os testes de inteligência aparecem usados como informação de suporte às opções dos alunos em face da multiplicidade de saídas académicas e profissionais, diferenciadas entre si também pelas exigências cognitivas que envolvem. Claro que os testes de inteligência podem também servir uma abordagem mais dinâmica quer das dificuldades de aprendizagem quer das opções vocacionais dos alunos. Por exemplo, os testes de inteligência podem constituir uma oportunidade de auto-conhecimento dos sujeitos em prol da exploração e definição de percursos escolares e profissionais futuros (Abreu, 2001; Almeida, 1996b; Leitão e Paixão, 1999). Ainda, é possível aceitar e acreditar que, não só as aptidões interferem nas aprendizagens escolares, como estas podem ter impacto no desenvolvimento e na diferenciação das próprias aptidões (Almeida, 1988b, 1996a; Almeida e Campos, 1986).

Com este artigo pretendemos apresentar novos dados sobre a validade de critério ou externa dos resultados numa bateria de provas de raciocínio em uso no nosso país. Ao fazê-lo pretendemos acompanhar e actualizar dados relativos ao funcionamento destas mesmas provas, assim como perceber se a evolução do sistema educativo tem mantido a tradicional relação entre habilidades cognitivas e resultados académicos dos alunos. Mais concretamente neste estudo, recorrendo à versão da Bateria de Provas de Raciocínio para os alunos do 5º e 6º anos (BPR5-6), pretendemos verificar que níveis de correlação tais provas apresentam com as classificações escolares e se tais coeficientes oscilam em função da natureza das provas (conteúdo dos respectivos itens) de acordo com alguma investigação anterior com este tipo de provas (Almeida, 1988a; Almeida e Martins, 1996; Almeida, Antunes, Martins e Primi, 1997; Meuris et al., 1988).

Método

Amostra

A amostra é composta por 1163 alunos, repartidos pelo 5º (n=552) e 6º (n=611) anos de escolaridade. Esta amostra foi constituída a nível nacional (Continente) na base de critérios de representatividade e aleatoriedade tomando a distribuição dos alunos por zonas do país, comunidades urbanas e rurais, ano de escolaridade e género, com base nos dados estatísticos obtidos junto do Ministério da Educação. As idades dos alunos oscilam entre os 9 e os 17 anos, com uma média de 11,2 e um desvio-padrão de 1,08.

Instrumento

Este estudo reporta-se à aferição nacional da Bateria de Provas de Raciocínio, versão para os alunos do 5º e 6º anos de escolaridade (BPR5-6; Almeida, 2003). Esta bateria é formada por 4 provas, com limite de tempo, todas elas apelando o raciocínio: Prova RA (20 itens figurativo-abstractos, no formato de analogia, com quatro alternativas de resposta), Prova RN (15 itens numéricos, no formato de sequência de números, lineares ou alternadas, que o sujeito deverá continuar), Prova RV (20 itens verbais, no formato de analogia de palavras, com quatro alternativas de resposta) e Prova RP (15 itens de resolução de problemas lógico-dedutivos, onde o sujeito tem que elaborar a resposta). A par das provas psicológicas, foram recolhidos alguns dados relativos ao percurso escolar dos alunos (número de reprovações prévias e classificações actuais em várias disciplinas).

Procedimento

A bateria foi aplicada colectivamente, em contexto de turma e de sala de aula, em tempos lectivos cedidos pelos professores. Os alunos foram informados dos contornos do estudo e da confidencialidade dos dados, bem como da importância da sua participação numa investigação deste género, reforçando o seu papel na representação de outros colegas do mesmo ano, da mesma escola e, numa perspectiva mais alargada, dos alunos da sua zona do País. O resultado de cada prova corresponde ao número de itens correctamente respondidos, sendo que no caso da Prova RN a cotação dos itens oscila de 0 a 2 (1 ponto quando o sujeito tiver respondido correctamente nos valores mas invertido a posição dos mesmos na sequência).

Resultados

No Quadro I apresentamos a análise descritiva das variáveis em estudo, nomeadamente no que toca o desempenho cognitivo dos alunos (as quatro provas de raciocínio e na bateria no seu todo, BPR5-6) e o seu rendimento escolar (em diversas disciplinas, em particular o Português, a Matemática, as Ciências da Natureza, a Educação Visual e Tecnológica, a Educação Física e uma nota global de rendimento académico, que reflecte, no essencial, a nota média dos resultados escolares dos alunos nas primeiras três disciplinas, ditas mais académicas - Nota), tomando o sexo e os dois anos de escolaridade.

Quadro I – Estatística descritiva das variáveis avaliadas

Variáveis		5º Ano				6º Ano			
		Mín.	Máx.	Méd.	DP	Mín.	Máx.	Méd.	DP
Masculino	RA	1.0	19.0	10.2	3.59	1.0	20.0	11.1	3.44
	RV	1.0	20.0	11.6	3.57	3.0	20.0	13.0	3.30
	RN	2.0	28.0	13.9	5.67	2.0	30.0	15.2	6.53
	RP	1.0	14.0	8.4	3.04	1.0	15.0	9.0	2.80
	BPR5-6	3.9	17.0	10.6	2.85	4.5	17.0	11.7	2.76
	Português	2.0	5.0	3.3	.74	1.0	5.0	2.9	.76
	Matemática	2.0	5.0	3.5	.86	2.0	5.0	3.0	.89
	C.Natureza	2.0	5.0	3.5	.85	2.0	5.0	3.2	.79
	E.V.T.	2.0	5.0	3.3	.63	2.0	5.0	3.2	.67
	E.Física	2.0	5.0	3.6	.64	2.0	5.0	3.5	.72
	Nota	2.0	5.0	3.4	.73	1.7	5.0	3.0	.72
Feminino	RA	1.0	17.0	10.2	3.28	1.0	19.0	11.0	3.27
	RV	3.0	20.0	11.7	3.24	4.0	20.0	13.0	3.30
	RN	2.0	26.0	11.2	5.42	2.0	28.0	14.0	5.57
	RP	2.0	20.0	9.0	2.44	1.0	15.0	9.6	2.34
	BPR5-6	4.2	17.1	10.4	2.62	3.4	18.1	11.6	2.57
	Português	2.0	5.0	3.3	.70	2.0	5.0	3.2	.82
	Matemática	2.0	5.0	3.4	.83	2.0	5.0	3.2	.92
	C.Natureza	2.0	5.0	3.4	.78	2.0	5.0	3.5	.81
	E.V.T.	2.0	5.0	3.5	.65	2.0	5.0	3.5	.70
	E.Física	2.0	5.0	3.4	.62	2.0	5.0	3.3	.53
	Nota	2.0	5.0	3.4	.67	2.0	5.0	3.3	.75

Como se poderia antecipar, os resultados nas provas de raciocínio aumentam, e de forma sistemática, quando passamos do 5º para o 6º ano de escolaridade (o que desde logo pode ser um sinal da validade dos resultados da bateria). Também podemos aceitar que os resultados dos alunos nas provas de raciocínio numérico e de resolução de problemas são mais baixos que nas provas de raciocínio verbal e abstracto. O conteúdo numérico e os cálculos necessários na Prova RN, a par das exigências de leitura e compreensão na Prova RP, poderão explicar as maiores dificuldades dos alunos na realização destas duas provas. Não emergem diferenças nas médias tomando os alunos de acordo com o sexo.

Em relação ao rendimento escolar dos alunos, as classificações nas diversas disciplinas consideradas ou no seu rendimento médio, oscilam maioritariamente entre 2 e 5, sendo a média das classificações nos dois anos escolares positiva (acima de 3.0). Tendencialmente as médias das classificações escolares são mais baixas nas disciplinas de Português e Matemática, situação mais clara junto dos alunos do 6º ano e do sexo masculino.

No quadro II descrevemos as correlações dos resultados dos alunos nessas disciplinas e as quatro provas psicológicas aplicadas, em função do ano de escolaridade que frequentam. A par das provas e disciplinas isoladas, toma-se na análise a média nas quatro provas da bateria (BPR5-6) e a média das classificações nas disciplinas de Português, Matemática e Ciências da Natureza (Nota).

Quadro II – Correlações entre os resultados nas provas psicológicas e as notas escolares em função do ano

Ano	Disciplinas	Prova RA	Prova RN	Prova RV	Prova RP	BPR5-6
5º	Português	.29**	.36***	.41***	.49***	.50***
	Matemática	.45***	.50***	.50***	.54***	.64***
	C. Natureza	.40***	.39***	.45***	.48***	.56***
	E.V.T.	.32***	.24**	.22*	.32***	.35***
	E. Física	.09	.19*	.13	.11	.17*
	Nota	.44***	.49***	.52***	.58***	.65***
6º	Português	.34***	.36***	.48***	.51***	.55***
	Matemática	.45***	.43***	.45***	.46***	.59***
	C. Natureza	.41***	.36***	.45***	.48***	.55***
	E.V.T.	.23*	.16*	.20*	.22**	.28**
	E. Física	.11	.12	.12	.11	.16*
	Nota	.45***	.43***	.52***	.54***	.63***

Two tailed test: * $p < .05$; ** $p < .01$; *** $p < .001$

Numa primeira análise, podemos verificar que as correlações entre uma medida global de habilidade cognitiva (BPR5-6) e os resultados escolares nas diversas disciplinas são consideráveis e mais ou menos equivalentes nos dois anos de escolaridade em análise. Observando em maior detalhe, encontramos dados particularmente interessantes aquando das correlações fortes entre os resultados obtidos pelos alunos na bateria e o seu desempenho nas disciplinas de cariz eminentemente académico, ou se quisermos, mais intelectuais (como o Português, a Matemática e as Ciências da Natureza), daí termos procedido ao cálculo de uma nota global de rendimento académico (nota média dos resultados escolares dos alunos nas três disciplinas supramencionadas) e confirmado as correlações bastantes elevadas que estabelecem com a nota global de habilidade cognitiva, quer no 5º ano ($r = .65$, $p < .001$), quer no 6º ano ($r = .63$, $p < .001$). As correlações são mais baixas entre as provas cognitivas e as disciplinas de Educação Visual e Tecnológica (EVT) e de Educação Física.

Interessante apontar as correlações mais elevadas obtidas pela prova de resolução de problemas (Prova RP), o que nos parece ficar a dever-se à necessidade de leitura, compreensão e interpretação de pequenos textos por parte dos alunos. Ao mesmo tempo, assiste-se a uma tendência para coeficientes de correlação mais elevados quando cruzamos disciplinas e provas mais próximas em termos de conteúdo. Esta situação está presente, por exemplo, quando cruzamos a Prova RV e o Português ou quando cruzamos a Prova RA e a Educação Visual e Tecnológica.

Uma outra forma de analisarmos a relação entre o desempenho dos alunos nas provas da BPR5-6 e o seu rendimento escolar passa pela análise de eventuais diferenças nas médias dos alunos nas provas de raciocínio em função do número de reprovações escolares em anos anteriores (constituição de quatro grupos: 0 reprovações, 1 reprovação, 2 reprovações e 3 ou mais reprovações). Para esta análise procedemos a uma análise da variância (*F-one-way* com procedimento *scheffe* para apreciação de contrastes entre os grupos). Dada a proximidade dos alunos do 5º e 6º ano, optámos por tomar a amostra de alunos na sua globalidade, sendo de acrescentar que alguns alunos não preencheram este campo do inquérito (*missing*). No quadro III apresentamos os resultados desta análise, podendo assumi-la como novo contributo ao estudo da validade externa da BPR5-6.

Quadro III – Resultados nas provas em função do nível de reprovações dos alunos

Prova	Nível de Reprovações	n	Méd.	DP	F	Prob.	Contraste								
RA	0	920	11.1	3.24	25.688	.000	0>1,2,3								
	1	142	9.1	3.40											
	2	59	8.7	3.91											
	3 ou mais	24	8.5	3.73											
RV	0	923	12.8	3.25	36.449										
	1	144	10.6	3.34											
	2	59	10.1	3.11											
	3 ou mais	24	9.5	3.13											
RN	0	910	14.4	5.92	28.039										
	1	138	10.1	5.00											
	2	55	10.9	5.03											
	3 ou mais	25	11.0	5.99											
RP	0	923	9.5	2.47	60.605										
	1	141	7.2	2.85											
	2	56	7.1	2.17											
	3 ou mais	24	5.8	2.82											
BPR5-6	0	907	11.6	2.61	54.155										
	1	134	9.1	2.43											
	2	52	9.1	2.22											
	3 ou mais	22	8.5	2.69											

Os resultados comprovam uma diferenciação com significado estatístico nos resultados das quatro provas e no seu conjunto tomando os alunos diferenciados com base no número de reprovações escolares anteriores. Melhores níveis de realização nas provas de raciocínio ocorrem junto dos alunos com menor taxa de reprovação, como seria esperado numa lógica de validade das provas. Assim, ao longo das provas assistimos a uma diminuição progressiva nas médias dos alunos quando passamos do grupo de alunos sem qualquer reprovação ou com menos reprovações para os grupos de alunos com mais reprovações anteriores (situação que apenas não ocorre na Prova RN). Os valores de F mostram-se estatisticamente significativos para as quatro provas individualizadas e para a média geral na bateria (BPR5-6), sendo essa diferença particularmente expressiva na Prova RP (o que aliás já ocorria cruzando os resultados nesta prova com as classificações nas disciplinas escolares).

Apreciando os contrastes nos resultados das quatro provas e no seu conjunto (BPR5-6) verifica-se um padrão comum de valores. Em termos de signifi-

cância estatística, apenas as comparações entre o grupo de alunos sem qualquer reprovação e qualquer um dos outros grupos de alunos se apresenta estatisticamente significativa. Aliás, olhando as médias dos resultados dos quatro grupos de alunos em comparação, podemos verificar que a maior diferença se estabelece entre alunos sem e com reprovações anteriores, independentemente do número dessas reprovações.

Conclusões

Tomando uma amostra nacional de alunos do 5º e 6º anos de escolaridade procurou-se neste artigo avaliar a validade por referência a critério da Bateria de Provas de Raciocínio (versão para alunos do 5º e 6º anos de escolaridade; BPR5-6). Como critério externo foi tomado o rendimento escolar dos alunos, repartido por duas informações: as classificações num conjunto de disciplinas e o número de reprovações anteriores. A maior utilização destas provas na prática psicológica em contextos escolares justifica o cuidado havido em verificar quão associado se encontra o desempenho nestas provas ao rendimento académico dos alunos (Almeida, 1996b).

As correlações obtidas, independentemente do ano escolar dos alunos, sugerem a validade da presente bateria, indo no mesmo sentido de estudos anteriores com versões próximas ou preliminares da actual BPR5-6 (Almeida, 1988a; Almeida e Martins, 1996; Almeida e Campos, 1986; Almeida et al., 1997). Em primeiro lugar, observam-se correlações positivas e estatisticamente significativas quando cruzamos desempenhos nos testes e as classificações escolares, em particular naquelas disciplinas mais tipicamente académicas (Português, Matemática, Ciências da Natureza). As correlações são mais baixas quando se reportam às disciplinas de Educação Visual e Tecnológica (EVT) e de Educação Física. Esta discrepância nos coeficientes de correlação consoante a tipologia de disciplinas escolares reforça a validade de critério da bateria.

Em segundo lugar, assistimos que os coeficientes de correlação oscilam em função das provas da bateria. A prova de resolução de problemas (Prova RP) e a prova de raciocínio verbal (Prova RV) apresentam coeficientes de correlação mais elevados com as classificações escolares, sugerindo a importância da linguagem no sucesso escolar dos alunos. Esta situação parece-nos particularmente presente nos itens da Prova RP pois que os mesmos envolvem algumas competências de leitura, compreensão e interpretação de enunciados dos problemas (pequenos textos).

Em terceiro lugar, e de acordo com alguma literatura na área, assistimos, já ao nível dos alunos do 5º e do 6º ano de escolaridade, a alguma tendência para os coeficientes de correlação serem mais elevados quando se reportam a provas de raciocínio e a disciplinas curriculares que se aproximam em termos de conteúdo (por exemplo Prova RV e a disciplina de Português). Do mesmo modo, os coeficientes de correlação são mais elevados quando cruzamos uma nota global nas quatro provas da bateria (BPR5-6) e uma nota global nas disciplinas escolares. A investigação sugere que estas notas globais acabam por melhor representar a diversidade de variáveis inerentes ao funcionamento e desempenho cognitivo (cf. Almeida, 1988a; Almeida e Campos, 1986), merecendo particular destaque o facto das correlações agora obtidas se aproximarem de .60 (ou seja a possibilidade de associarmos 35% da variância nos resultados escolares destes alunos ao seu desempenho na BPR5-6).

Finalmente, os resultados nas quatro provas e no seu conjunto apresentam-se diferenciados em função dos índices de reprovação escolar dos alunos. A análise de variância mostrou que os alunos sem qualquer reprovação escolar apresentam médias superiores na bateria, sendo essa diferença estatisticamente significativa. Os valores obtidos apontam que a diferença se situa comparando os alunos com e sem reprovação, não parecendo relevante para tais diferenças o ter uma, duas ou mais reprovações. Como ocorreu em relação às classificações nas várias disciplinas, também neste critério externo de validade dos resultados da BPR5-6, a prova de resolução de problemas (Prova RP) mostra-se a mais associada ao índice de reprovação dos alunos em anos anteriores, sugerindo a sua relevância na análise das dificuldades escolares dos alunos, uma das situações que mais justificam o trabalho dos psicólogos nas escolas.

Referências

- Abreu, M. V. (2001). Desenvolvimento vocacional e estratégias de motivação para aprendizagens consistentes. *Psychologica*, 26, 9-26.
- Almeida, L. S. (1988a). *O raciocínio diferencial de jovens*. Porto: INIC.
- Almeida, L. S. (1988b). O impacto das experiências educativas na diferenciação cognitiva dos alunos: Análise dos resultados em provas de raciocínio diferencial. *Revista Portuguesa de Psicologia*, 24, 131-157.
- Almeida, L. S. (1992). Inteligência e aprendizagem: Dos seus relacionamentos à sua promoção. *Psicologia: Teoria e Pesquisa*, 8, 272-292.
- Almeida, L. S. (1994). *Inteligência: Definição e medida*. Aveiro: CIDInE.
- Almeida, L. S. (1996a). Cognição e aprendizagem: Como a sua aproximação conceptual pode favorecer o desempenho cognitivo e a realização escolar. *Psicologia: Teoria, Investigação e Prática*, 1, 17-32.
- Almeida, L. S. (1996b). O espaço das aptidões cognitivas e dos respectivos testes na investigação e prática psicológica. *Revista Portuguesa de Pedagogia*, XXX (2), 71-81.
- Almeida, L. S. (2003). *Bateria de Provas de Raciocínio: Versão para o 5º e 6º anos de escolaridade (BPR5-6)*. Braga: Universidade do Minho, Instituto de Educação e Psicologia.
- Almeida, L. S. e Campos, B. P. (1986). Validade preditiva dos testes de raciocínio diferencial. *Cadernos de Consulta Psicológica*, 2, 105-118.
- Almeida, L. S. e Martins, T. (1996). A validade de critério em provas cognitivas: Contributos para um banco de dados na população portuguesa. In *Actas da IV Conferência Internacional "Avaliação Psicológica: Formas e Contextos"* (pp.577-584). Braga: Associação dos Psicólogos Portugueses.
- Almeida, L. S., Antunes, A. M., Martins, T. B. O. e Primi, R. (1997). Bateria de Provas de Raciocínio (BPR-5): Estudo de validação em contexto escolar. In *Actas do III Congresso Galaico-Português de Psicopedagogia* (pp. 355-365). Corunha: Universidade da Corunha.
- Bleichrodt, N., Drenth, P. J. D., Zaal, J. N. e Resing, W. C. M. (1984). *Revision Amsterdam Child Intelligence Test*. Lisse, The Netherlands: Swets.
- Bleichrodt, N., Resing, W. C. M., Drenth, P. J. D. e Zaal, J. N. (1987). *The measurement of children's intelligence*. Lisse, The Netherlands: Sweets.
- Jensen, A. R. (1980). *Bias in mental testing*. London: Methuen.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.
- Leitão, L. M. e Paixão, A. P. (1999). Contributos para um modelo integrado de orientação escolar e profissional no Ensino Superior. *Psicologia: Teoria, Investigação e Prática*, 4 (1), 191-209.

- Meuris, G., Almeida, L. S. e Campos, B. P. (1988). Le rôle de l'opération mentale et du contenu des items dans la performance cognitive. *Bulletin de Psychologie Scolaire et d'Orientation*, 37 (1), 12-21.
- Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J., e Urbina, S. (1996). Intelligence, knowns, and unknowns. *American Psychologist*, 51, 77-101.
- Pereira, M., & Simões, M. R. (2005, no prelo). A avaliação da inteligência nas dificuldades de aprendizagem: Investigação com a WISC-III. *Psicologia, Educação e Cultura*, 9(2) [neste número]
- Primi R. e Almeida, L. S. (2000). Estudo de validação da bateria de provas de raciocínio (BPR-5). *Psicologia: Teoria e Pesquisa*, 16 (2), 165-173.
- Schmidt, F. L. e Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262-274.
- Schmidt, F. L., Pearlman, K. e Hunter, J. E. (1980). The validity and fairness of employment and educational tests for Hispanic Americans: A review and analysis. *Personnel Psychology*, 33, 705-724.
- Simões, M. R. & Albuquerque, C. P. (2002). Estudos com a versão portuguesa da WISC-III no âmbito de validade preditiva: Relação com as classificações escolares. *Psychologica*, 29, 153-168.
- Simões, M. R., Gomes, A. A., Albuquerque, C. P., Almeida, L. S., Pereira, M., Santos, M. J. & Fonseca, A. C. (1998). A avaliação da inteligência nas dificuldades de aprendizagem: Estudos com a WISC-III. In L. S. Almeida, M. J. Gomes, P. B. Albuquerque & S. G. Caires (Eds.), *Actas do IV Congresso Galaico-Português de Psicopedagogia* (pp. 462-473). Braga: Universidade do Minho.
- Simões, M. R., Seabra-Santos, M. J., Albuquerque, C. P., Pereira, M. M., Almeida, L. S., Ferreira, C., Lopes, A. F., Gomes, A. A., Xavier, R. E., Rodrigues, F., Lança, C., Barros, J., San Juan, L., & Oliveira, E. (2003). Escala de Inteligência de Wechsler para Crianças – Terceira Edição (WISC-III). In M. M. Gonçalves, M. R. Simões, L. S. Almeida & C. Machado (Coords.), *Avaliação Psicológica: Instrumentos validados para a população portuguesa* (Vol. I, pp. 221-252). Coimbra: Quarteto.
- Snow, R. E. e Yalow, E. (1982). Education and intelligence. In Robert J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 493-585). Cambridge: Cambridge University Press.
- Sternberg, R. J. e Kaufman, J. C. (1996). Innovation in intelligence testing: The curious case of the dog that didn't bark. *European Journal of Psychological Assessment*, 12 (3), 175-182.
- Te Nijenhuis, J., Evers, A. e Mur, J. (2000). The validity of the Differential Aptitude Test for the assessment of immigrant children. *Educational Psychology*, 20, 99-115.
- Te Nijenhuis, J., Tolboom, E. R. e Bleichrodt, N (2004). Does cultural background influence the intellectual performance of children from immigrant groups?: The RAKIT Intelligence Test for Immigrant Children. *European Journal of Psychological Assessment*, 20, 10-26.

COGNITIVE APTITUDES AND ACADEMIC ACHIEVEMENT: THE PREDICTIVE VALIDITY OF INTELLIGENCE TESTS

Leandro S. Almeida e Gina Lemos

Departamento de Psicologia, Instituto de Educação e Psicologia, Universidade do Minho

Abstract: Taking a Portuguese sample of students from 5th and 6th grades ($n=1163$) we analyse the validity of scores on “Bateria de Provas de Raciocínio” (BPR5-6; Almeida, 2003) on basis of students’ academic achievement (school marks and previous academic retention). In general, data suggest a significant relationship between academic performance and results on the four reasoning tests of battery (abstract reasoning, verbal reasoning, numerical reasoning, and problem-solving). Students with better academic performance present better scores on BPR5-6 tests, which means external criterion referenced validity. Scores on problem-solving and verbal reasoning subtests present higher correlation coefficients with academic achievement, as well as these coefficients are higher when we consider the school marks on more academic disciplines (Portuguese, Mathematics, and Nature Sciences). Also, higher correlations occur when we consider a global result on four reasoning tests and on school disciplines, for example $r=.65$ and $r=.63$ for the students on 5th and 6th grades, respectively. Finally, the results on cognitive tests are related to the number of students’ retention. Students without academic retentions present higher means on BPR5-6 tests. These data support the external validity of BPR5-6 and its utilization by school psychologists on students’ learning difficulties.

KEY-WORDS: *Intelligence, School achievement, Predictive validity, Reasoning tests, Psychological Assessment.*

VALIDADE PREDITIVA DA WPPSI-R EM RELAÇÃO AO DESEMPENHO ESCOLAR AO LONGO DO 1º CICLO

Maria João Seabra-Santos*

*Serviço de Avaliação Psicológica,
Faculdade de Psicologia e de Ciências da Educação, Universidade de Coimbra, Portugal*

Resumo

A validade preditiva da Escala de Inteligência de Wechsler para a Idade Pré-Escolar e Primária – Forma Revista (WPPSI-R), recentemente aferida para a população portuguesa (Wechsler, 2003), foi estudada a partir da aplicação da Escala a uma amostra de 37 crianças de 5 anos, antes da entrada na escola, e subsequente seguimento do desempenho escolar destas crianças ao longo do seu percurso pelo 1º ciclo. As correlações entre os QIs da WPPSI-R (realização, verbal e da escala completa) e os resultados escolares são significativas nos três níveis escolares estudados (1º, 2º e 4º anos), situando-se os respectivos valores entre .40 e .60, para 34 dos 39 coeficientes calculados. A utilidade preditiva da escala é superior para as disciplinas de Matemática e Estudo do Meio, em comparação com a Língua Portuguesa.

PALAVRAS-CHAVE: *WPPSI-R, validade preditiva, idade pré-escolar, rendimento escolar.*

Introdução

A Escala de Inteligência de Wechsler para a Idade Pré-Escolar e Primária – Forma Revista (WPPSI-R) foi publicada nos Estados Unidos em 1989 e posteriormente aferida e editada em Portugal (Wechsler, 2003). O processo de validação desta escala para a população portuguesa contemplou a realização de diversos estudos, nomeadamente: estudos de análise factorial (Wechsler, 2003); estudos de correlação com outros instrumentos, como a WISC-III, as Matrizes Progressivas Coloridas de Raven (Wechsler, 2003), a Escala de Desenvolvimento de Griffiths (Albuquerque, Seabra-Santos e Lança, 2001a), provas piagetianas (Seabra-Santos, no prelo) e um questionário relativo ao de-

Morada (address): Maria João Seabra Santos. Serviço de Avaliação Psicológica. Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra, Rua do Colégio Novo, 3001-802 Coimbra, Portugal. Tel.: 239851450; Fax: 239851462; e-mail: seabramj@fpce.uc.pt.

* Este texto teve o apoio do Projecto de Investigação "Adaptação e aferição portuguesas da WISC-III e da WPPSI-R" (financiado pela Fundação para a Ciência e a Tecnologia) e do Centro de Psicopedagogia da Universidade de Coimbra.

envolvimento, respondido pelos professores (Seabra-Santos, 1999); e estudos com grupos especiais, designadamente crianças com dificuldades escolares (Seabra-Santos, 1998), em risco ambiental (Albuquerque, Seabra-Santos e Lança, 2001b), com inteligência superior (Wechsler, 2003) e candidatas a antecipação da escolaridade obrigatória (Wechsler, 2003). Estes diversos estudos contribuíram para a caracterização da WPPSI-R enquanto instrumento potencialmente útil em contextos clínicos, educativos ou de investigação, susceptível de fornecer uma medida válida da inteligência de crianças em idade pré-escolar. Porém, atendendo à publicação recente da WPPSI-R portuguesa, nenhum estudo foi ainda publicado sobre a validade preditiva desta aferição.

Quando falamos em validade preditiva de um instrumento de avaliação referimo-nos a um tipo de validade empírica, ou referenciada a um critério externo, que tem a ver com a capacidade do instrumento para prever o desempenho individual noutras actividades, com interesse para a avaliação em causa (Anastasi, 1990). O termo “preditivo” aplica-se à antecipação de um critério localizado no futuro.

Os estudos de validade preditiva da Escala de Inteligência de Wechsler para a Idade Pré-Escolar e Primária, na sua edição original, pretenderam esclarecer a capacidade da escala para antecipar, quer resultados em testes de inteligência aplicados mais tarde, quer o desempenho escolar de um aluno, sendo estas últimas em número bastante superior.

Dada a escassez de estudos de validade preditiva relativos à forma revista (WPPSI-R, Wechsler, 1989) e atendendo à grande sobreposição verificada entre as duas versões (a de 1989 e a de 1967), apresentam-se no Quadro 1 os estudos de validade preditiva que visam a utilidade da forma original da WPPSI na previsão do êxito escolar. As provas utilizadas como critério, nas diversas pesquisas, são testes de desempenho/realização escolar (*school achievement*) que, nos primeiros anos de escolaridade, incidem fundamentalmente sobre competências de matemática, leitura e escrita. Tal como se pode observar, as investigações variam marcadamente quanto às amostras utilizadas, metodologia seguida e resultados encontrados. As populações representadas vão desde crianças da classe média (por exemplo, Kaufman, 1973; White e Jacobs, 1979) a sujeitos de nível socio-económico baixo, nalguns casos pertencentes a minorias étnicas e frequentando o Programa *Head Start* (Crockett, Rardin e Pasewark, 1976; Lowe et al., 1987). A maioria dos estudos incide sobre amostras de crianças americanas, mas um deles recorre a um grupo de sujeitos israelitas (Lieblich e Shinar, 1975), enquanto que um outro utiliza um grupo de crianças britânicas (Yule, Gold e Busch, 1982) e, outro ainda, canadianas (White e Jacobs, 1979). As

pesquisas divergem, também, quanto ao tipo de teste de desempenho escolar a que recorrem, sendo alguns de administração individual (Feshbach, Adelman e Fuller, 1977; White e Jacobs, 1979), e outros de aplicação colectiva (Lieblich e Shinar, 1975; Yule et al., 1982). O intervalo decorrido entre a aplicação da WPPSI e a obtenção do critério varia entre 4 meses (Kaufman, 1973) e 11 anos (Lowe et al., 1987; Yule et al., 1982).

Quadro 1 - Estudos de validade preditiva da WPPSI relativamente ao desempenho escolar

Estudo ⁽¹⁾	Amostra ⁽²⁾	Intervalo	Teste de desempenho escolar (nível escolar)	Correlações
Krebs (1969)	70 crianças no último ano pré-escolar (35 de NSE elevado/35 de NSE baixo)	1 ano	Stanford Achievement Test e Gilmore Oral Reading Paragraphs (1º ano)	Leitura x (NSE elevado) $Q_{L_R} : .35$ $Q_{L_V} : .32$ $Q_{L_{EC}} : .40$ Leitura x (NSE baixo) $Q_{L_R} : .61$ $Q_{L_V} : .59$ $Q_{L_{EC}} : .66$
Kaufman (1973)	31 crianças brancas com 6 anos, classe média	4 meses	Metropolitan Achievement Test (1º ano)	Leitura x $Q_{L_{EC}} : .37$ Matemática x $Q_{L_{EC}} : .30$ (n.s.) x $Q_{L_R}/Q_{L_V} : .30$ (n.s.)
Lieblich & Shinar (1975)	54 crianças israelitas com 5-6 anos, 1º ano de escolaridade, classe média	18 meses	Testes israelitas de Leitura e Aritmética (2º ano)	Leitura x $Q_{L_R} : .61$ $Q_{L_V} : .57$ $Q_{L_{EC}} : .63$ Aritmética x $Q_{L_R} : .73$ $Q_{L_V} : .64$ $Q_{L_{EC}} : .73$
Crockett, Rardin & Pasewark (1976)	35 crianças, média de idades: 5A;07m (Programa <i>Head Start</i>)	3-4 anos	Metropolitan Achievement Test	Leitura x $Q_{L_R} : .23$ (n.s.) $Q_{L_V} : .03$ (n.s.) $Q_{L_{EC}} : .17$ (n.s.) Matemática x $Q_{L_R} : .52$ $Q_{L_V} : .24$ (n.s.) $Q_{L_{EC}} : .43$
Feshbach, Adelman & Fuller (1977)	2 grupos (403+364 crianças), último ano pré-escolar, classe média	1, 2 e 3 anos	Cooperative Primary Reading Tests e equivalentes-grau (1º, 2º e 3º anos)	Leitura x $Q_{L_{EC}}$ 1º ano: .32 e .42; 2º ano: .40 e .48; 3º ano: .45 e .45
White & Jacobs (1979)	28 crianças canadianas com 3-5 anos, classe média	1 a 3 anos	Gray Oral Reading Test (1º ano)	Leitura x $Q_{L_R} : .51$ $Q_{L_V} : .54$ $Q_{L_{EC}} : .58$
Yule, Gold & Busch (1982)	84 crianças britânicas com 5 anos	11 anos	Sentence Reading Test (11º ano)	Leitura x $Q_{L_R} : .53$ $Q_{L_V} : .61$ $Q_{L_{EC}} : .61$
"	"	"	Vernon's Graded Word Spelling Test (11º ano)	Ortografia x $Q_{L_R} : .51$ $Q_{L_V} : .48$ $Q_{L_{EC}} : .53$
"	"	"	Vernon's Graded Arithmetic-Mathematics Test (11º ano)	Aritmética x $Q_{L_R} : .69$ $Q_{L_V} : .65$ $Q_{L_{EC}} : .72$

(cont.)

Estudo ⁽¹⁾	Amostra ⁽²⁾	Intervalo	Teste de desempenho escolar (nível escolar)	Correlações
Lowe, Anderson, Williams & Currie (1987)	amostra A: 169 crianças negras com 5-6 anos (Programa <i>Head Start</i>)	1 até 11 anos	Notas escolares (compósito) do 1º ao 11º ano de escolaridade	Q _R : .45-.64 Q _V : .31-.55 Q _{EC} : .45-.64
"	159 crianças de A	5 meses	Metropolitan Readiness Test (1º ano)	Q _R : .61 Q _V : .56 Q _{EC} : .52
"	"	"	Wide Range Achievement Test (1º ano)	Q _R : .40-.57 Q _V : .39-.57 Q _{EC} : .43-.62
"	101 crianças de A	5 anos	Iowa Tests of Basic Skills (5º ano)	Q _R : .43 Q _V : .37 Q _{EC} : .45
"	69 crianças de A	8 anos	California Achievement Test (8º ano)	Q _R : .24-.56 Q _V : .25-.35 Q _{EC} : .27-.52
"	72 crianças de A	11 anos	Iowa Tests of Educational Development (11º ano)	Q _R : .30 Q _V : .24 Q _{EC} : .30
McEvoy & Johnson (1989)	59 crianças mexicanas-americanas com 5 anos, NSE baixo	1 a 4 anos	Notas escolares em Leitura	Leitura x Q _R : .32 Q _V : .43 Q _{EC} : .47
"	"	"	Iowa Tests of Basic Skills (1º a 4º ano)	Leitura x Q _R : .36 Q _V : .63 Q _{EC} : .60

⁽¹⁾ Os estudos são apresentados por ordem cronológica

⁽²⁾ A idade das crianças refere-se ao momento da primeira avaliação

Embora se encontrem resultados algo divergentes entre os vários estudos, os coeficientes de correlação obtidos situam-se, na sua maioria, entre .40 e .60, indiciando uma capacidade moderada da WPPSI para prever o rendimento escolar futuro.

As inconsistências verificadas entre os diversos estudos sobre validade preditiva poderão, pelo menos em parte, estar relacionadas com as diferentes características dos sujeitos avaliados. O *nível socio-económico* poderá constituir uma variável a considerar, neste contexto. Assim, Krebs (1969) encontra em crianças pertencentes a meios desfavorecidos uma correlação de .66 entre o Q_{EC} e as competências de Leitura, superior à alcançada na mesma investigação mas para um grupo de crianças de nível socio-económico elevado (.40). Também McEvoy e Johnson (1989) registam um coeficiente de .60 entre as competências de Leitura e o Q_{EC}, igualmente em crianças de nível socio-

económico baixo. Já Crockett e cols. (1976), pelo contrário, chegam a coeficientes não significativos, também com a Leitura, e a uma correlação entre o QI_{EC} e a Matemática igual a .43, sempre em crianças de meios desfavorecidos. O género dos examinandos pode, igualmente, condicionar a capacidade preditiva da WPPSI. Neste sentido, Feshbach e cols. (1977) verificam uma maior possibilidade de antecipar o rendimento na Leitura e na Matemática no caso dos rapazes do que das raparigas, dado este que corrobora resultados similares registados no estudo de Crockett (1974), mas somente para a Matemática. Note-se, finalmente, que as duas investigações que registam correlações superiores (Lieblich e Shinar, 1975; Yule et al., 1982) recorrem a amostras de *crianças não americanas* e a versões da Escala adaptadas às respectivas populações (aferição israelita e adaptação britânica, respectivamente). Este conjunto de resultados poderá indicar que a capacidade da WPPSI para prever o rendimento escolar futuro é variável, consoante as características da população em causa.

Bastante encorajadores relativamente à validade preditiva da WPPSI são os valores registados por Yule e cols. (1982), que mostram ser o desempenho na WPPSI capaz de prever 51% da variância dos resultados escolares alcançados passados 11 anos. Aliás, analisadas as várias pesquisas no seu conjunto, não se observa uma diminuição dos coeficientes de correlação, à medida que aumenta o intervalo de tempo entre as duas avaliações. Veja-se, a este propósito, que o estudo em que se verificam correlações mais baixas (.30 a .37) recorre a um intervalo de somente 4 meses (Kaufman, 1973), ao passo que na pesquisa de Yule e cols., apesar da intromissão de um lapso de 11 anos, se obtêm correlações entre .50 e .70. No estudo de Lowe e cols. (1987), o cálculo de correlações entre os resultados da WPPSI-R e um índice geral de desempenho escolar revela, igualmente, valores estáveis ao longo de 11 anos de escolaridade, sendo o valor mais elevado (.65) registado no 9º ano.

Apesar das diferenças observadas entre os estudos quanto à magnitude dos coeficientes de correlação, registam-se, igualmente, alguns pontos de convergência entre eles. Assim, verifica-se que a validade da WPPSI tende a ser maior na antecipação do desempenho em Aritmética do que na previsão do rendimento noutras disciplinas escolares (Crockett et al., 1976; Lieblich e Shinar, 1975; Yule et al., 1982). Por outro lado, os subtestes da Escala de Wechsler que são referenciados como mais preditivos, num maior número de estudos, são os de Aritmética, Vocabulário e Figuras Geométricas (Crockett et al., 1976; Krebs, 1969; Segerstrom, 1976; Yule et al., 1982). E embora a subescala verbal contenha um maior número de conteúdos relacio-

nados com a escolaridade do que a de realização, o QI_V surge como mais preditivo do que o QI_R somente em alguns dos estudos inventariados no Quadro 1 (por exemplo McEvoy e Johnson, 1989), mas não na generalidade deles.

Quadro 2 - Estudos de validade preditiva da WPPSI-R relativamente ao desempenho escolar

Estudo ⁽¹⁾	Amostra ⁽²⁾	Intervalo	Teste de desempenho escolar (nível escolar)	Correlações
Kaplan (1993)	50 crianças com 4-6 anos, classe média ou média-alta	2 anos	Comprehensive Testing Program II (1º ano)	QI_R : .13-.35(n.s.) QI_V : .44-.71 QI_{EC} : .38-.65
Kaplan (1996)	91 crianças com 4-6 anos, classe média ou média-alta	2, 3 e 4 anos	Comprehensive Testing Program III (1º, 2º e 3º anos)	QI_R : (n.s.) 1º Ano: QI_V : .41-.65 QI_{EC} : .42-.50 2º Ano: QI_V : .54-.61 QI_{EC} : .48-.60 3º Ano: QI_V : .45-.62 QI_{EC} : .28-.58

(1) Os estudos são apresentados por ordem cronológica

(2) A idade das crianças refere-se ao momento da primeira avaliação. O segundo estudo inclui parte da amostra sobre a qual incidiu o primeiro.

No Quadro 2 encontram-se sintetizados elementos relativos a duas pesquisas sobre a validade preditiva da versão revista da WPPSI (Kaplan, 1993, 1996), as duas únicas disponíveis até à data, tanto quanto é do nosso conhecimento. Estes estudos põem em evidência o elevado valor preditivo de resultados obtidos na WPPSI-R, antes do último ano pré-escolar (*pre-kindergarten*), relativamente ao rendimento escolar no final do 1º, 2º e 3º anos de escolaridade, em crianças da classe média ou média-alta, com inteligência média ou acima da média. Nestas condições, o QI_V e os subtestes verbais mostram-se particularmente preditivos, enquanto que o QI_R e os subtestes de realização surgem como preditores negligenciáveis, discrepância esta que não havia sido verificada no conjunto de estudos incidindo sobre a versão inicial da WPPSI. Os coeficientes de correlação referentes aos subtestes de realização, sendo baixos e quase todos não significativos, são responsáveis pela inferioridade das correlações registadas com o QI_{EC} , relativamente às obtidas com o QI_V . Os subtestes de Informação, Aritmética, Vocabulário e Semelhanças, surgem como os mais preditivos – os de Aritmética e Vocabulário não se destacam dos outros dois quanto a este aspecto, contrariamente ao que acontecia nos estudos com a primeira versão da WPPSI.

Observa-se, igualmente, que a utilidade preditiva da WPPSI-R em relação ao desempenho em Matemática, no final do 1º ano de escolaridade, não é superior à relativa às restantes matérias. Porém, no final do 3º ano, o teste de Matemática é, justamente, aquele com o qual se registam correlações mais elevadas (.62 para o QI_V), à semelhança do que se havia verificado com a forma original da WPPSI. Por outro lado, também para a WPPSI-R se verifica que o alongamento do intervalo de tempo, entre a avaliação com a Escala e a obtenção do critério, não é acompanhado por uma diminuição da relação entre ambos, pelo menos dentro de um período de quatro anos (Kaplan, 1996).

Esta revisão da literatura aponta algumas pistas quanto à validade preditiva da WPPSI, ao mesmo tempo que deixa várias questões em aberto. Uma ideia fundamental a reter é que as divergências entre os resultados obtidos nos diversos estudos reflectem diferenças na validade preditiva da WPPSI, quando usada com objectivos distintos ou com populações de características diversas.

No estudo que passamos a apresentar acompanhámos o desempenho escolar, ao longo do 1º ciclo, de uma amostra de crianças que haviam sido avaliadas com a WPPSI-R no seu último ano pré-escolar. Trata-se, pois, de um estudo de tipo longitudinal, através do qual pretendemos contribuir para a caracterização da validade preditiva da WPPSI-R, na sua versão portuguesa.

Metodologia

Amostra

A presente investigação assenta numa amostra de crianças residentes no concelho de Coimbra, avaliadas com a WPPSI-R no ano anterior à sua entrada no 1º ano do ensino básico e seguidas ao longo do 1º ciclo. A aplicação da WPPSI-R teve lugar no contexto de um conjunto de estudos relativos à validade desta escala de inteligência na população portuguesa, os quais se basearam numa amostra constituída por 120 crianças. A selecção desta amostra partiu de uma tiragem aleatória de jardins-de-infância do concelho de Coimbra, tendo, igualmente, sido seleccionadas algumas escolas do 1º ciclo, pela proximidade geográfica que apresentavam relativamente aos jardins-de-infância tirados aleatoriamente. A selecção das crianças dentro de cada estabelecimento escolar foi também feita de modo aleatório, de entre aquelas que satisfaziam as exigências da amostra estratificada em termos de idade, género e localidade de residência (zona urbana *versus* pequena localidade do concelho).

Para o presente estudo foram retidos todos os sujeitos da amostra total que iriam ingressar no 1º ciclo do ensino básico no ano lectivo subsequente, o que representou um total de 37 crianças. O desempenho escolar destas 37 crianças foi acompanhado ao longo do 1º ciclo de escolaridade, tendo sido possível obter dados relativos a 34 dessas crianças no final do 1º e do 2º anos, e de 33 no final do 4º, o que representa uma taxa de atrito bastante baixa, correspondente às crianças que não foi possível localizar.

Das 34 crianças seguidas pelo menos durante os dois primeiros anos de escolaridade, 18 eram raparigas e 16 eram rapazes; 18 residiam em zona urbana e 16 em pequenas localidades do concelho; as suas idades estavam compreendidas entre os 5 anos e 4 meses e os 5 anos e 11 meses (média igual a 5 anos e 7 meses); o número médio de anos de escolaridade frequentado pelas respectivas mães situava-se nos 8 anos (mínimo 4, máximo 16); eram maioritariamente provenientes de famílias de nível socio-económico baixo ou médio; no momento em que foram avaliadas com a WPPSI-R, as crianças frequentavam jardins-de-infância (9, no total) dos quais alguns da rede pública (para 13 das crianças) e outros particulares (para 21 das crianças), sendo a maioria destes últimos Instituições Particulares de Solidariedade Social (IPSS).

Procedimento

As crianças foram avaliadas com a WPPSI-R no último trimestre pré-escolar. O rendimento escolar destas crianças foi seguido ao longo do 1º ciclo, tendo sido obtidos dados relativos ao seu desempenho escolar em 3 momentos: no final do 1º, do 2º e do 4º anos de escolaridade, isto é, respectivamente um, dois e quatro anos após a aplicação da WPPSI-R.

Instrumentos

Escala de Inteligência de Wechsler para a Idade Pré-Escolar e Primária — Revista (WPPSI-R): A Escala de Inteligência de Wechsler para a Idade Pré-Escolar e Primária — Forma Revista (Wechsler, 1989) é um instrumento administrado individualmente, que avalia a inteligência de crianças pequenas. Trata-se de um dos poucos instrumentos de avaliação psicológica, destinados à idade pré-escolar, que se encontram aferidos para a população portuguesa, possuindo normas para crianças com idades compreendidas entre os 3 anos e os 6 anos e 6 meses (Wechsler, 2003). Tal como as outras escalas de inteligência de Wechsler, a WPPSI-R é um instrumento compósito,

com duas subescalas: uma subescala de Realização, que compreende os subtestes de Composição de Objectos, Figuras Geométricas, Quadrados, Labirintos e Completamento de Gravuras, e a partir da qual é possível calcular um QI de Realização (QI_R); e uma subescala Verbal, que inclui os subtestes de Informação, Compreensão, Aritmética, Vocabulário e Semelhanças, e com base na qual se calcula um QI Verbal (QI_V). O QI da Escala Completa (QI_{EC}) é obtido a partir dos resultados dos 10 subtestes. Os resultados padronizados para os subtestes têm média 10 e desvio-padrão 3, enquanto que os valores correspondentes para os QIs são, respectivamente, iguais a 100 e a 15.

Avaliação do desempenho escolar: O desempenho escolar foi avaliado com base nas apreciações dos professores relativas às áreas curriculares de Matemática, Estudo do Meio e Língua Portuguesa (no primeiro ano esta última disciplina foi desdobrada em duas vertentes: Leitura e Escrita). No 1º ano de escolaridade a apreciação do desempenho dos alunos foi feita com base num registo de avaliação estandardizado, adoptado por algumas escolas. Neste, o rendimento dos alunos em cada área curricular era classificado em 4 categorias (insuficiente; suficiente; bom; muito bom), as quais foram posteriormente convertidas numa classificação quantitativa numa escala de 1 a 4. Contudo, este sistema de recolha de dados não se revelou eficaz, dada a variabilidade dos sistemas de avaliação adoptados pelos diversos professores nas várias instituições. Por este motivo, nos anos subsequentes, optou-se por pedir a todos os professores que classificassem os alunos numa escala de 1 a 5, em cada uma das áreas referidas.

Resultados e Discussão

As médias e desvios-padrão dos resultados obtidos na WPSI-R são apresentados no Quadro 3. Tal como se pode observar, as crianças que compõem a amostra em estudo funcionam, em termos médios, a um nível próximo da amostra normativa (média do QI_{EC} igual a 98.91, com um mínimo de 76 e um máximo de 123). Por outro lado, a análise dos desvios-padrão permite caracterizar esta amostra como bastante homogénea no que toca aos resultados obtidos nos subtestes verbais (desvio-padrão do QI_V igual a 8.92, contra o valor 15 da amostra de aferição).

Quadro 3 - Médias e desvios-padrão dos subtestes e escalas da WPPSI-R

Subteste/Escala	Média	d.p.
Composição Objectos	9.03	3.42
Figuras Gemétricas	10.50	3.15
Quadrados	9.41	2.66
Labirintos	10.29	2.02
Completamento Gravuras	9.21	2.33
Tabuleiro Animais	10.41	3.29
Informação	10.18	1.85
Compreensão	10.82	2.24
Aritmética	9.85	2.09
Vocabulário	10.85	1.89
Semelhanças	10.26	2.06
Frases Memorizadas	9.79	1.77
QI _R	96.21	14.36
QI _V	102.76	8.92
QI _{EC}	98.91	11.82

A constatação desta excessiva homogeneidade, que tende a baixar os coeficientes de correlação, levou a que se optasse por introduzir uma correcção, no cálculo destes coeficientes, por forma a obter uma estimativa das correlações para a população total, cujos desvios-padrão de cada subteste e das escalas de QI são, respectivamente, iguais a 3 e a 15⁽¹⁾. Os coeficientes de correlação assim corrigidos são apresentados no Quadro 4. Dado o elevado número de coeficientes de correlação calculados optou-se por adoptar critérios de significância estatística exigentes (.01 e .005), no sentido de diminuir a possibilidade de obter coeficientes significativos devido ao acaso. Ressalte-se o facto de os coeficientes de correlação relativos ao 1º ano não serem rigorosamente comparáveis aos calculados para os anos seguintes, podendo ser um pouco mais baixos, uma vez que nesse 1º ano os professores classificaram as crianças numa escala de 4 pontos, ao invés de 5, como aconteceu nos dois níveis subsequentes.

⁽¹⁾ Aplicou-se, para este efeito, a fórmula proposta por Guilford e Fruchter (1978, p.325). Esta fórmula parte do valor do coeficiente de correlação obtido numa amostra de amplitude restrita, como é o caso da amostra do presente estudo, e conhecendo o valor dos desvios-padrão na amostra com restrição de amplitude e numa amostra sem restrição de amplitude (neste caso, a amostra normativa), estima o valor do coeficiente de correlação que seria obtido na amostra sem restrição de amplitude.

Quadro 4 - Correlação entre a WPSI-R e o desempenho escolar no 1º, 2º e 4º anos⁽¹⁾

Resultados escolares													
WPPSI-R	1º ano ⁽²⁾					2º ano ⁽²⁾				4º ano ⁽²⁾			
	E. Meio	Leitura	Escrita	Matemt.	Média 1º	E. Meio	L. Port	Matemt.	Média 2º	E. Meio	L. Port.	Matemt.	Média 2º
Composição Objectos	.38	.27	.35	.38	.39*	.38	.32	.29	.35	.36	.32	.31	.35
Figuras Geométricas	.43**	.51**	.56**	.48**	.56**	.41*	.32	.33	.37	.34	.35	.45**	.40*
Quadrados	.58**	.30	.48**	.50**	.52**	.39*	.29	.31	.35	.46**	.36	.46**	.45**
Labirintos	.06	-.04	-.16	.18	.00	.00	-.09	.01	-.03	.03	-.19	.00	-.05
Completamento Gravuras	.41*	.28	.31	.53**	.42**	.58**	.50**	.45**	.53**	.44**	.43**	.48**	.48**
Tabuleiro Animais	.28	.10	.05	.29	.19	-.01	.02	.11	.04	.14	.07	.19	.15
Informação	.46**	.49**	.22	.46**	.45**	.33	.18	.08	.20	.30	.11	.37	.28
Compreensão	.31	.05	.07	.26	.19	.32	.15	.18	.23	.30	-.05	.21	.17
Aritmética	.44**	.39*	.32	.46**	.45**	.19	.17	.40*	.27	.36	.16	.52**	.38
Vocabulário	.51**	.38	.40*	.41*	.47**	.54**	.36	.17	.38	.40*	.41*	.26	.38
Semelhanças	.19	.19	.34	.03	.22	.30	.40	.27	.34	.30	.33	.30	.33
Frases Memorizadas	.08	.23	.06	.01	.11	-.03	-.12	-.08	-.08	-.06	.08	-.03	-.01
QI _R	.52**	.39*	.47**	.56**	.54**	.48**	.38*	.38	.43**	.46**	.40*	.48**	.47**
QI _V	.62**	.48**	.45**	.53**	.57**	.56**	.43**	.39*	.48**	.55**	.32	.55**	.51**
QI _{EC}	.61**	.48**	.53**	.62**	.62**	.58**	.46**	.45**	.52**	.57**	.44**	.58**	.56**

* p<.01 **p<.005

⁽¹⁾ Coeficientes de correlação corrigidos em função da restrição de amplitude dos resultados da WPSI-R⁽²⁾ Estudo do Meio, Língua Portuguesa (Leitura, Escrita), Matemática, Média das diversas disciplinas.

A análise dos coeficientes de correlação apresentados no Quadro 4 permite verificar que os resultados obtidos na WPSI-R antes da entrada na escola são preditivos do êxito escolar, mesmo com um desfasamento temporal de 4 anos. A apreciação das correlações entre subtestes individuais e os resultados escolares mostra que alguns deles se revelam melhores preditores do que outros. Os subtestes em que se registam coeficientes estatisticamente significativos são, na subescala de realização, o Completamento de Gravuras, as Figuras Geométricas e os Quadrados; e na subescala verbal, o Vocabulário, a Informação e a Aritmética. Porém, enquanto que os coeficientes significativos observados em relação aos subtestes de realização se mantêm relativamente estáveis ao longo da escolaridade, os coeficientes obtidos para os subtestes verbais tendem a diminuir, sendo poucos os que se conservam acima do limiar de significância estatística, no final do 1º ciclo.

Os coeficientes mais elevados obtidos no presente estudo, no que toca aos subtestes individuais⁽²⁾, foram registados:

⁽²⁾ As indicações das aptidões avaliadas pelos vários subtestes foram retiradas de Kaufman e Lichtenberger (2000).

- *No 1º ano de escolaridade:* entre o subteste Figuras Geométricas – que mede a organização visomotora, síntese e processamento visual – e a Leitura e Escrita (.51 e .56, respectivamente); entre o subteste Quadradados – que avalia a visualização espacial, inteligência fluida e aptidão de conceptualização abstracta – e a disciplina de Estudo do Meio (.58); entre o subteste Vocabulário – que envolve o desenvolvimento da linguagem, formação de conceitos, memória e quantidade de informação – e igualmente a disciplina de Estudo do Meio (.51); e entre o subteste Completamento de Gravuras – que avalia a atenção aos detalhes, processamento simultâneo e memória visual a longo prazo – e a disciplina de Matemática;
- *No 2º ano de escolaridade:* entre os subtestes Completamento de Gravuras e Vocabulário e a disciplina de Estudo do Meio (.58 e .54, respectivamente).
- *No 4º ano de escolaridade:* entre o subteste Aritmética – que avalia o raciocínio numérico, cálculo mental, memória e conceitos quantitativos – e a disciplina de Matemática (.52).

Na presente investigação não se verifica a acentuada superioridade assinalada por Kaplan (1993, 1996) das correlações obtidas para os subtestes e subescala verbais, relativamente às obtidas para os subtestes e subescala de realização. Ao contrário, os coeficientes alcançados para ambas as subescalas apresentam valores próximos, o que é mais consonante com alguns dos estudos que incidiram sobre a primeira versão da WPPSI (Lieblich e Shinar, 1975; Low et al., 1987; Yule et al., 1982).

As correlações entre os QIs da WPPSI-R (realização, verbal e da escala completa) e os resultados escolares são significativas nos três níveis escolares estudados (1º, 2º e 4º anos), situando-se os respectivos valores entre .40 e .60 para 34 dos 39 coeficientes calculados. Podemos, pois, afirmar que os coeficientes de correlação obtidos no presente estudo são bastante elevados em comparação com os obtidos nas investigações publicadas sobre a validade preditiva da WPPSI ou WPPSI-R, sobretudo se tivermos em conta que, na generalidade destas investigações, o critério externo considerado – resultados obtidos em testes estandardizados de desempenho escolar – é mais objectivo que o utilizado na presente pesquisa na qual, dada a ausência de instrumentos estandardizados de avaliação das aprendizagens escolares, se recorreu a apreciações subjectivas feitas pelos professores sobre o rendimento académico dos seus alunos.

O valor preditivo da WPPSI-R é superior relativamente às disciplinas de Estudo do Meio e Matemática do que à de Língua Portuguesa, excepção feita para o 2º ano, em que os valores dos coeficientes relativos à Língua Portuguesa são, em geral, ligeiramente superiores aos registados com a Matemática.

Finalmente, e pese embora a ressalva feita antes quanto à possibilidade de os valores dos coeficientes obtidos para o 1º ano estarem um pouco deflacionados, o aumento do intervalo de tempo entre a aplicação da WPPSI-R e a obtenção do critério não parece ter um efeito corrosivo ao nível dos coeficientes de correlação verificando-se, pelo contrário, uma acentuada estabilidade. Isto significa que o valor preditivo da WPPSI-R se mantém ao longo do tempo, pelo menos durante o 1º ciclo de escolaridade. Esta verificação é, aliás, coadjuvada pelos resultados de vários dos estudos acima analisados (Feshbach et al., 1977; Kaplan, 1993, 1996; Lowe et al., 1987; Yule et al. 1982).

A possibilidade de generalizar os resultados obtidos na presente investigação deve ser encarada com precaução, dada a dimensão relativamente reduzida da amostra utilizada. A replicação do presente estudo com amostras de características diversas e de maiores dimensões poderá, no futuro, dar lugar a conclusões mais finas sobre a validade preditiva da WPPSI-R. Noutras investigações a realizar será interessante ter em conta, nomeadamente, algumas variáveis susceptíveis de influenciar a utilidade preditiva da Escala, como parece ser o caso do género dos sujeitos (Crockett, 1974; Feshbach et al., 1977).

Para além do mais, parece-nos plausível a hipótese de que a possibilidade de antecipar o desempenho escolar dos alunos pode ser aumentada se os resultados obtidos na WPPSI-R forem utilizados em conjugação com outros tipos de informação, designadamente dados obtidos a partir de questionários preenchidos pelos professores, ou características demográficas das crianças tais como a proveniência socio-económica ou as habilitações dos pais. Contamos, num futuro próximo, poder apresentar resultados empíricos que comprovem esta hipótese.

Referências

- Albuquerque, C. P., Seabra-Santos, M. J. e Lança, C. (2001a). Avaliação da inteligência na idade pré-escolar: Relações entre a WPPSI-R e a Escala de Desenvolvimento de Griffiths. In B. D. Silva e L. S. Almeida (Eds.), *Actas do VI Congresso Galaico-Português de Psicopedagogia* (vol. 2, pp. 747-758). Braga: Centro de Estudos em Educação e Psicologia.
- Albuquerque, C. P., Seabra-Santos, M. J. e Lança, C. (2001b). Análise do desempenho de crianças em risco ambiental na WPPSI-R (Wechsler Preschool and Primary Scale of Intelligence – Revised). *Revista Galego-Portuguesa de Psicoloxía e Educación*, 5, 193-202.
- Anastasi, A. (1990). *Psychological testing* (6th ed.). New York: Macmillan.
- Crockett, B. K., Rardin, M. W. e Pasewark, R. A. (1976). Relationship of WPPSI and subsequent Metropolitan Achievement Test scores in Head-Start children. *Psychology in the Schools*, 13, 19-20.
- Feshbach, S., Adelman, H. e Fuller, W. (1977). Prediction of reading and related academic problems. *Journal of Educational Psychology*, 69, 299-308.
- Guilford, J. P. e Fruchter, B. (1978). *Fundamental statistics in psychology and education* (6th ed.). New York: McGraw-Hill.
- Kaplan, C. (1993). Predicting first-grade achievement from pre-kindergarten WPPSI-R scores. *Journal of Psychoeducational Assessment*, 11, 133-138.
- Kaplan, C. (1996). Predictive validity of the WPPSI-R: A four year follow-up study. *Psychology in the Schools*, 33, 211-220.
- Kaufman, A. S. (1973). Comparison of the WPPSI, Stanford-Binet, and McCarthy Scales as predictors of first-grade achievement. *Perceptual and Motor Skills*, 36, 67-73.
- Kaufman, A. S. e Lichtenberger, E. O. (2000). *Essentials of WISC-III and WPPSI-R assessment*. New York: John Wiley & Sons.
- Krebs, E. G. (1969). The Wechsler Preschool and Primary Scale of Intelligence and prediction of reading achievement in first grade. *Dissertation Abstracts International*, 30, 4279A.
- Lieblich, A. e Shinar, M. (1975). The predictive validity of the WPPSI with Israeli children. *Educational and Psychological Measurement*, 35, 473-475.
- Lowe, J. D., Anderson, H. N., Williams, A. e Currie, B. B. (1987). Long-term predictive validity of the WPPSI and the WISC-R with Black school children. *Personality and Individual Differences*, 8, 551-559.
- McEvoy, R. E. e Johnson, D. L. (1989). Comparison of an intelligence test and a screening battery as predictors of reading ability in low income, Mexican American children. *Hispanic Journal of Behavioral Sciences*, 11, 274-282.

- Seabra-Santos, M. J. (1998). *WPPSI-R: Estudos de adaptação e validação em crianças portuguesas*. Tese de doutoramento apresentada à Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra.
- Seabra-Santos, M. J. (1999). Avaliação da inteligência em idade pré-escolar e escolar: Relação com as apreciações dos professores. *Revista Portuguesa de Pedagogia*, 33 (2), 5-22.
- Seabra-Santos, M. J. (no prelo). Avaliação da inteligência: Relação entre o desempenho na WPPSI-R e em provas piagetianas. *Psicologia e Educação*.
- Segerstrom, R. R. (1976). A study in the prediction of school readiness and school failure. *Dissertation Abstracts International*, 7050A.
- Wechsler, D. (1967). *Manual for the Wechsler Preschool and Primary Scale of Intelligence*. New York: Psychological Corporation.
- Wechsler, D. (1989). *WPPSI-R Manual*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (2003). *Escala de inteligência de Wechsler para a idade pré-escolar e primária – Edição revista*. Lisboa: CEGOC. (Aferição para a população portuguesa de Seabra-Santos, M. J., Simões, M. R., Rocha, A. M. e Ferreira, C.)
- White, D. R. e Jacobs, E. (1979). The prediction of first-grade reading achievement from WPPSI scores of preschool children. *Psychology in the Schools*, 16, 189-192.
- Yule, W., Gold, R. D. e Busch, C. (1982). Long-term predictive validity of the WPPSI: An 11-year follow-up study. *Personality and Individual Differences*, 3, 65-71.

PREDICTIVE VALIDITY OF THE WPPSI-R IN RELATION TO ACADEMIC ACHIEVEMENT DURING THE FIRST CYCLE

Maria João Seabra-Santos

Faculdade de Psicologia e de Ciências da Educação, Universidade de Coimbra, Portugal

Abstract: The predictive validity of the Wechsler Intelligence Scale for Children – Revised (WPPSI-R), recently standardized in Portugal (Wechsler, 2003), was studied using academic achievement as a criterion. Thirty-seven children were assessed with the WPPSI-R when they were five-years-old and their academic achievement was followed during the first cycle. Correlations between IQs (performance, verbal and full scale) and scores of academic performance endorsed by teachers at the end of grades 1, 2 and 4 are significant, with values ranging from .40 to .60 to 34 of the 39 coefficients calculated. The predictive utility of the WPPSI-R is higher for some subjects (Maths and Science) than for others (Portuguese).

KEY-WORDS: *WPPSI-R, predictive validity, preschool children, school achievement.*

A AVALIAÇÃO DA INTELIGÊNCIA NAS DIFICULDADES DE APRENDIZAGEM: INVESTIGAÇÃO COM A WISC-III

Marcelino Pereira

Mário R. Simões*

Serviço de Avaliação Psicológica, Faculdade de Psicologia e de Ciências da Educação, Universidade de Coimbra

Resumo

Nesta investigação, o nosso objectivo é fornecer um quadro de referência para a interpretação dos resultados alcançados na WISC-III (Escala de Inteligência de Wechsler para Crianças - Terceira Edição) por crianças com dificuldades de aprendizagem. Para o efeito comparamos os desempenhos de uma amostra de 54 crianças que os seus professores sinalizaram como apresentando D.A., com os resultados de um grupo de controlo, recorrendo à metodologia de *pares idênticos*. São analisadas as seguintes variáveis: (i) distribuição dos resultados nos treze testes que compõem a escala; (ii) hierarquização dos resultados pelos três índices da solução factorial (Compreensão Verbal, Organização Perceptiva e Velocidade de Processamento); (iii) resultados em duas medidas de dispersão (valor da discrepância entre o QI Verbal e o QI de Realização; distribuição dos treze testes em torno da média individual – *scatter*).

PALAVRAS-CHAVE: *WISC-III; Avaliação da inteligência; Dificuldades de aprendizagem.*

Introdução

A avaliação da inteligência no contexto das necessidades educativas especiais, impõe-se de forma natural desde os seus primórdios, se atendermos ao facto de que a primeira escala de inteligência (Escala Métrica de Inteligência, Binet, 1905) foi criada com esse intuito: identificar as crianças que teriam dificuldades no seu percurso escolar de modo a lhes proporcionar um ensino diferenciado.

Morada (address): Marcelino Pereira. Serviço de Avaliação Psicológica. Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra, Rua do Colégio Novo, 3001-802 Coimbra, Portugal. Tel.: 239851450; Fax: 239851462; e-mail: marcelinopereira@sapo.pt.

* Este texto teve o apoio dos Projectos de Investigação "Adaptação e aferição portuguesas da WISC-III e da WPPSI-R" e "Adaptação e Standardização Portuguesa de Testes Neuropsicológicos" (POCTI/PSI/35410/2000), financiados pela Fundação para a Ciência e Tecnologia e pelo Centro de Psicopedagogia da Universidade de Coimbra.

No âmbito específico das dificuldades de aprendizagem (D.A.), as escalas de inteligência assumem ainda maior protagonismo, porquanto constituem o elemento chave para o diagnóstico deste problema. De facto, na definição mais consensual das dificuldades de aprendizagem, o critério de discrepância entre a aptidão (potencial medido através de uma escala de inteligência) e o desempenho (resultados escolares), é de longe o mais prevalente (e.g., Truscott e Frank, 2001). Ou seja, de acordo com este critério, o que caracteriza todas as crianças com dificuldades de aprendizagem é a disparidade entre os seus resultados escolares (obtidos num teste de desempenho académico) e os que legitimamente se poderiam esperar, tendo em conta a sua idade e nível intelectual. O critério da discrepância pode ser obtido recorrendo a diferentes procedimentos metodológicos. Actualmente, a solução mais frequente compara o desempenho alcançado numa escala de inteligência (valor do QI) com o resultado obtido num teste aferido de conhecimentos escolares. Em países onde existem adaptações destes dois tipos de provas, considera-se que o diferencial entre elas tem significado clínico quando iguala ou ultrapassa o valor de 1.5 desvio-padrão. Por exemplo, se a criança apresenta um QI de 109 e apenas alcança um resultado de 87 num teste aferido de conhecimentos escolares. Refira-se que o valor definido para estabelecer o ponto-de-corte é variável: assim nalguns casos impõe-se um diferencial de 2 desvios-padrão, enquanto noutros, a diferença de 16 pontos é critério suficiente. Numa tentativa de reduzir essa arbitrariedade, já é frequente recorrer a modelos de regressão que permitem definir com maior rigor os pontos-de-corte. Com base nesta metodologia, nos Estados Unidos, considera-se que a discrepância média entre o QI da Escala Completa e o desempenho académico em testes standardizados é de 17.3 para a leitura, 16.28 para a matemática e 22.85 na linguagem escrita (Kaufman e Lichtenberger, 1999). Refere-se ainda, como critério de diagnóstico, que as dificuldades de aprendizagem podem ocorrer conjuntamente com outros problemas (por exemplo, perturbações sensoriais), mas não devem ser consideradas como o resultado directo desses *handicaps*; por último, as D.A. poderão ser generalizadas ou circunscritas a uma determinada área (DSM-IV; *American Psychiatric Association*, 1994).

Saliente-se que a relativa inoperância dos critérios supra-referidos, tem levado muitos autores a atribuir maior importância ao grau da dificuldade de modo a se poder diferenciar com mais precisão as D.A. de outros problemas de aprendizagem, cuja ocorrência, etiologia e gravidade são distintas (Gersons-Wolfensberger e Ruijsnaars, 1997). A mesma orientação é retomada em 2001 pela *Learning Disabilities Association of Ontario* (LDAO).

Por último, refira-se que a caracterização semiológica dos distúrbios de aprendizagem é variável nas diferentes faixas etárias e, em rigor, o diagnóstico só se poderá fazer nas crianças em idade escolar. Quer isto dizer, que, embora alguns sintomas ou percursos dos distúrbios de aprendizagem possam estar presentes na idade pré-escolar – devendo por isso merecer a nossa atenção, sobretudo numa perspectiva de acção profiláctica – eles não podem, contudo, operar como critério de diagnóstico (Rourke, 1991).

Saliente-se que, de acordo com as estatísticas de 1999 do Departamento Federal de Educação nos Estados Unidos, as crianças com dificuldades de aprendizagem constituem mais de 50% da população que recebe apoios educativos especiais e que as escalas de inteligência são aplicadas anualmente a um milhão de crianças, pois os seus resultados são decisivos para a admissão aos programas de educação especial (Watkins e Kush, 2002). Deixando transparecer a importância do uso das Escalas de Inteligência de Wechsler nos grupos especiais da população, nomeadamente como ferramenta imprescindível para o diagnóstico das dificuldades de aprendizagem, a *Psychological Corporation* editou, em 1992, o WIAT (*Wechsler Individual Achievement Test*). O teste é constituído por oito provas (competências básicas de leitura, raciocínio matemático, ortografia, compreensão da leitura, operações numéricas, compreensão auditiva, expressão oral e expressão escrita), cujos valores compósitos se distribuem por quatro áreas: leitura, matemática, linguagem e escrita. Nos casos das dificuldades de aprendizagem, recomenda-se que a WISC-III e o WIAT sejam usados em conjunto de modo a permitir avaliar a discrepância entre o QI e o desempenho (Flanagan e Alfonso, 1993). No sentido de operacionalizar o critério de discrepância, o manual do WIAT apresenta uma tabela com os pontos-de-corte a partir dos quais diferenças entre as duas medidas (WISC-III e WIAT) assumem significado clínico.

Sendo inquestionável que as escalas de inteligência são um instrumento imprescindível no protocolo de avaliação psicológica das crianças com dificuldades de aprendizagem, é, no entanto, mais polémica a definição do ponto-de-corte a partir do qual é legítimo formular esse diagnóstico. Numa abordagem clássica, o valor mínimo de QI a alcançar fixa-se em 90. Mais recentemente, alguns autores (e.g., Siegel, 1989; Vidal e Manjón, 2000) propõem uma descida desse valor para 80, alegando que o desempenho nas escalas de inteligência, nomeadamente nas Escalas de Inteligência de Wechsler, envolvem competências (como por exemplo, a linguagem expressiva, a velocidade de processamento da informação e a atenção dirigida) que, por definição, estão comprometidas nas crianças com D.A. e, consequentemente, pro-

duzem enviesamentos na avaliação do seu potencial intelectual, subestimando-o. Neste contexto, é necessário encontrar mecanismos que minimizem o impacto da interferência desses factores, e a redução do limite mínimo do QI tem sido a fórmula mais aplicada.

Podemos encontrar uma outra forma de atenuar essas limitações no uso de metodologias alternativas que analisam a distribuição dos resultados nos diferentes testes da escala, não se cingindo à leitura dos valores de QI. Ou seja, assume-se que os resultados disponibilizados nas Escalas de Inteligência de Wechsler poderão apresentar outras potencialidades interpretativas, sendo encarados como uma medida de funções cognitivas mais específicas que ultrapassam a simples mensuração do factor g (identificado a partir do QI da Escala Completa) permitindo, por isso, o estudo das diferenças individuais em domínios particulares da inteligência. Em simultâneo, na prática clínica, essas análises poderão actuar como “marcadores” ou sinalizadores de um determinado quadro nosográfico, constituindo uma prova adicional para a fundamentação do diagnóstico (Kaufman, 1994; Prifitera, Weiss e Saklofske, 1998).

Neste âmbito, em contexto clínico, têm sido desenvolvidos sistemas interpretativos com o objectivo de identificar padrões específicos de testes ou perfis (*ranking* dos resultados obtidos nos diferentes testes da escala pelo mesmo sujeito) que permitam formular hipóteses sobre as áreas fortes e fracas de um determinado indivíduo, orientando intervenções casuísticas diferenciadas. Assume-se assim, pelo menos implicitamente, que um determinado perfil nos diferentes testes da escala é preditivo de uma *performance* específica nos domínios académico, social ou outro (Bolen, Aichinger, Hall e Webster, 1995; Webster, 1988; citados por Carlton e Sapp, 1997). Por exemplo, a presença do perfil ACID (Resultados mais baixos nos testes de Aritmética, Código, Informação e Memória de Dígitos) estaria presente nos casos de dificuldades de aprendizagem e nos problemas de comportamento, nomeadamente na hiperactividade (e.g., Reynolds e Kaufman, 1990; Sandoval, 1994; Prifitera e Dresh, 1993).

De acordo com Watkins, Kush e Glutting (1997), um dos percursos desta prática foi o próprio Wechsler, quando em 1958 avançou com a hipótese de que a esquizofrenia infantil poderia ser diagnosticada através da presença de um perfil típico na WISC, caracterizado por resultados elevados no Completamento de Gravuras e Composição de Objectos, acompanhados por baixos desempenhos na Disposição de Gravuras e Memória de Dígitos. Outro pioneiro da análise de perfis é Bannatyne, ao sugerir, em 1968, que a análise da distribuição dos resultados dos testes nas escalas de Wechsler, poderia ser-

vir para identificar as crianças com distúrbios de aprendizagem, nomeadamente os casos de “dislexia genética”. Esta autora sugere a presença de quatro categorias de factores: (i) *Conceptual*, que agrupa os testes de Semelhanças, Vocabulário e Compreensão; (ii) *Espacial*, que inclui os testes de Completamento de Gravuras, Cubos e Composição de Objectos; (iii) *Sequencial*, constituído pelos testes de Aritmética, Memória de Dígitos e Código; (iv) *Conhecimento Adquirido*, que engloba os testes de Informação, Vocabulário e Aritmética.

As recategorizações em perfis de desempenho

A análise dos resultados nas Escalas de Inteligência de Wechsler, através do agrupamento dos diferentes testes em unidades com maior significado clínico e utilidade diagnóstica, tem constituído uma linha de investigação permanente nas sucessivas revisões das escalas. Neste âmbito, o perfil mais conhecido e popular, particularmente nos casos de dificuldades de aprendizagem, é o factor Resistência à Distracção (constituído pelos testes de Aritmética, Código e Memória de Dígitos), isolado pela primeira vez na análise factorial da WISC-R operada por Kaufman em 1975. Originalmente este “factor” foi referenciado como uma medida de memória auditiva imediata e seria capaz de diferenciar as crianças com distúrbios de aprendizagem das crianças com desempenho escolar regular. Todavia, o seu poder discriminativo nesse campo revelou-se diminuto e, talvez por isso, os investigadores adicionaram ao trio inicial o subteste da Informação, criando o perfil ACID.

Teoricamente a ocorrência do perfil ACID deverá ser mais frequente nas crianças com distúrbios de aprendizagem ou outros problemas da infância, tais como a hiperactividade (e.g., Watkins, Kush e Glutting, 1997; Kaufman e Leichstenberger, 1999; Reynolds e Kaufman, 1990; Sandoval, 1994; Prifitera e Dersh, 1993). Nesse sentido é comum o recurso a estudos comparativos, em que o perfil de desempenho encontrado em amostras clínicas (dificuldades de aprendizagem) é comparado com o padrão de resultados observado na amostra de aferição ou em grupos de controlo. Como veremos já a seguir, os dados estão longe de reunir consenso.

Watkins, Kush e Glutting (1997) referem diversas investigações cujos resultados atestam a falácia do factor resistência à distração e do factor ACID como medidas credíveis de diagnóstico (cf. Barkley, DuPaul e McMurray, 1990; Gussin e Javorsky, 1995; Kavale e Forness, 1984; McDermott, Fantuzzo e Glutting, 1990).

Os resultados da pesquisa conduzida pelos próprios autores (Watkins, Kush e Glutting, op. cit.), com o objectivo de analisar a validade preditiva e discriminante do perfil ACID, também não constituem prova abonatória para o uso deste tipo de resultados. Na sua amostra, de 612 crianças com dificuldades de aprendizagem, embora constatem que o perfil ACID (que ocorre quando os quatro testes que o formam são os mais baixos da escala) é mais frequente na amostra das crianças com dificuldades de aprendizagem (4.1% versus 1.1%, no grupo de controlo) ele é, no entanto, um critério de diagnóstico manifestamente insuficiente gerando um grande número de falsos negativos. Isto é, 587 (96%) das 612 crianças com distúrbios de aprendizagem foram incorrectamente classificadas como não pertencendo a esta categoria. Paralelamente, esse mesmo critério gerou 1% de falsos positivos ou seja, 24 crianças sem dificuldades de aprendizagem foram incorrectamente incluídas no grupo D.A. Por seu turno, a precisão acerca da presença do distúrbio (verdadeiros positivos), foi de apenas 4%, identificando correctamente 25 crianças numa amostra de 612. Acresce ainda, que o perfil ACID também não é um preditor robusto do desempenho académico da criança com D.A. Isto é, os resultados alcançados pelos sujeitos nos testes de leitura, expressão escrita e matemática não estão associados ao perfil ACID.

Golombok e Rust (1992) num estudo levado a cabo no âmbito da aferição inglesa da WISC-III, também não observam, de forma pronunciada, a presença do factor ACID num grupo de crianças com distúrbios de aprendizagem moderados e severos. Assim, referem que no grupo com dificuldades de aprendizagem apenas 5.1% dos sujeitos registam a totalidade do factor (ou seja, os resultados nos testes de Aritmética, Código, Informação e Dígitos, são iguais ou inferiores ao resultado mais baixo alcançado nos restantes testes da escala, excluindo os Labirintos e a Pesquisa de Símbolos), enquanto na amostra de aferição a prevalência é de 1.1%.

Por seu turno, Prifitera e Dersh (1993) constataam que o perfil ACID surge mais frequentemente em grupos especiais da população do que na amostra de aferição. Assim, ele está presente em 5% dos casos de distúrbios de aprendizagem e em 12% das crianças com distúrbio de défice de atenção com hiperactividade, enquanto na amostra de aferição a sua ocorrência é de 1%. Face a estes resultados, Prifitera e Dersh concluem que a presença de um determinado padrão de desempenho (perfil) na WISC sugere a existência de um determinado distúrbio. A mesma posição é assumida por Sattler (1988, cit. por Prifitera e Dersh, 1993). Este autor, num estudo de revisão que envolveu a análise de diversas investigações, seriou por ordem decrescente os resultados dos doze testes da WISC-R obtidos pelas crianças com dificuldades

de aprendizagem. Verificou que os testes que compõem o perfil ACID são os mais fracassados e conclui que há evidência suficiente para afirmar que a sua presença nas crianças com dificuldades de aprendizagem é um dado adquirido.

Por sua vez, Ward, Ward, Hatt, Young, e Mollner (1995), num total de 719 alunos, em que 165 apresentavam dificuldades de aprendizagem, referem que o perfil ACID tem maior prevalência no grupo das dificuldades de aprendizagem. No entanto, a incidência observada, apenas 4.7%, é insuficiente para validar esse perfil como marcador fiável das D.A..

A inoperância do perfil ACID como critério de diagnóstico das dificuldades de aprendizagem, também emerge quando as amostras são constituídas por adultos. Por exemplo, Johnson e Blalock (1987) observam que numa amostra de 93 adultos com distúrbios de aprendizagem, apenas 11 casos apresentam resultados inferiores nos quatro testes que constituem o perfil ACID. Ainda na população adulta, Kaufman e Leichtenberger (1999) constataam que numa amostra de 46 indivíduos apenas 6.5% apresentava o perfil ACID. Embora se trate de um valor superior ao registado na população geral, é manifestamente insuficiente para ser tomado como característica de grupo e, mais uma vez, o perfil ACID falha nas projecções que a partir dele se poderão fazer acerca do funcionamento cognitivo e/ou comportamental do indivíduo.

Resumindo, os dados da investigação mostram-nos que o perfil ACID não pode assumir-se como um marcador fiável do diagnóstico clínico das crianças com D.A., mas pode revelar-se de alguma utilidade nos processos referentes à tomada de decisão e à planificação das intervenções.

Mais recentemente, com a inclusão do teste de Pesquisa de Símbolos na WISC-III, Prifitera e Dersch (1993) combinaram os testes dos índices Resistência à Distracção (Aritmética e Memória de Dígitos) e Velocidade de Processamento (Pesquisa de Símbolos e Código), criando um novo perfil, que Kaufman (1994) baptizou com a designação de SCAD. Este perfil tem sido encarado por alguns autores como uma alternativa mais credível do que o ACID para sinalizar as crianças com dificuldades de aprendizagem, sobretudo quando a sua presença está associada a resultados elevados no factor Organização Perceptiva (Prifitera e Dersch, 1993; Kaufman e Lichtenberger, 2000; Ward et al., 1995). No entanto, os resultados de investigações mais recentes não fornecem suporte a esta hipótese. Neste plano, Watkins e Glutting (2000) elencam um conjunto de investigações que não valida este novo perfil de resultados (cf. Dumont, Farr, Willis e Whelley, 1998; Lipsitz, Dworkin e Erlenmeyer-Kimling, 1993; McDermott e Glutting, 1997; Rispen *et al.*, 1997).

Os resultados da investigação levada a cabo por Watkins e Glutting (2000) também não são encorajadores. Estes autores compararam os desempenhos de dois grupos de crianças em testes de leitura e matemática, um com D.A. e outro sem esse problema. Em simultâneo, analisaram os perfis que essas crianças apresentavam na WISC-III, de modo a poderem testar a eficácia preditiva dessa medida nos desempenhos em provas de leitura e matemática. Constataram que o perfil dos resultados na escala de Wechsler explicava apenas 4.8% da variância dos resultados obtidos no teste de leitura e 9.1% da variância dos resultados alcançados no teste de matemática. Watkins e Glutting concluem que a análise dos perfis nas escalas de inteligência de Wechsler apresenta uma fraca validade preditiva do funcionamento cognitivo do sujeito, e que, mesmo assim, estão provavelmente inflacionados pelo erro de medida.

Uma outra proposta de análise da distribuição dos resultados alcançados na WISC pelas crianças com D.A., é feita por Mishra, Lord e Sabers (1989), tendo como suporte teórico o modelo de processamento da informação de Luria (*Luria-Das Model of Successive and Simultaneous Processing*). De acordo com Luria, 1971 (citado por Mishra, Lord e Sabers, 1989), o funcionamento cognitivo é o produto de uma interação dinâmica, na qual a componente codificação da informação é um elemento crucial que recorre a dois tipos de analisadores: o processamento sucessivo e o processamento simultâneo. Mishra, Lord e Sabers (1989) defendem que os testes de Memória de Dígitos e Código são uma medida válida da componente tratamento sucessivo da informação, enquanto o Completamento de Gravuras, Cubos e Composição de Objectos são os que melhor representam o modo de tratamento simultâneo da informação. Com base nestes postulados, prevêem que a análise destes padrões de desempenho possa ser útil nos casos do diagnóstico das D.A.. Todavia, desconhecemos investigações que tenham testado este modelo de análise.

Em face dos resultados até agora alcançados, Kaufman e Lichtenberger (2000) concluem, prudentemente, que a análise dos diferentes padrões de desempenho nas Escalas de Inteligência de Wechsler, nomeadamente na WISC-III, embora apareçam associados de forma consistente a determinados grupos clínicos, não são suficientemente “fortes” para operarem ao nível do diagnóstico diferencial. Ou seja, a presença ou ausência de um padrão específico de desempenho, por exemplo o ACID ou o SCAD, constitui uma peça preciosa na análise interpretativa dos resultados, fornecendo-nos informação acerca das aptidões cognitivas da criança, mas não legitima uma tomada de decisão diagnóstica. Mais do que a presença ou ausência de um determina-

do perfil, a história de desenvolvimento, a história familiar, a história médica, a observação de comportamentos e a análise de protocolos escolares da criança assumem-se como medidas mais fiáveis para clarificar um diagnóstico de D.A..

Assumindo uma posição mais definitiva, encontramos Mayes, Calhoun e Crowell, 1998 (citados por Kaufman e Lichtenberger, 2000), ao referirem que os perfis das crianças com D.A. na WISC-III não se diferenciam dos restantes.

Por último, gostaríamos de salientar que à relativa diversidade dos resultados alcançados nas diferentes investigações, provavelmente, não serão estranhos os problemas que se levantam na constituição das amostras de crianças com D.A., que operam com critérios de referência distintos, inviabilizando, frequentemente as comparações.

Índices ou factores

A estrutura factorial da WISC-R compreende três factores que abarcam diferentes testes: Compreensão Verbal (Informação, Semelhanças, Vocabulário e Compreensão); Organização Perceptiva (Completamento de Gravuras; Disposição de Gravuras, Cubos e Composição de Objectos) e Resistência à Distracção (Aritmética, Memória de Dígitos e Código). Por seu turno, a análise factorial da WISC-III, derivada a partir da amostra de aferição americana, aponta para uma solução de quatro factores: os dois primeiros rigorosamente iguais aos da WISC-R; o factor Resistência à Distracção inclui apenas os testes de Aritmética e Memória de Dígitos; e, surge um quarto factor, Velocidade de Processamento, que integra o teste de Código e um novo teste de realização, suplementar, denominado Pesquisa de Símbolos (Wechsler, 1991). Curiosamente, este último teste tinha sido introduzido com o objectivo de clarificar a natureza do terceiro factor da WISC-R.

Posteriormente, análises factoriais independentes (levadas a cabo com outros dados que não os da aferição americana) como as de Sattler, 1992, e Thorndike, 1992, ambos citados por Kush (1996), revelam que a estrutura factorial que melhor descreve os dados da WISC-III é a que corresponde a uma solução de três factores. Por outro lado, os dados de investigações posteriores também começam a sugerir que a Velocidade de Processamento é o elemento que melhor caracteriza o terceiro factor da WISC-III. É neste sentido que apontam igualmente os estudos de análise factorial com a versão portuguesa da WISC-III (Wechsler, 2003). Uma solução de três factores é aquela que parece estar mais ajustada aos dados portugueses: Compreensão Verbal

(Informação, Semelhanças, Vocabulário e Compreensão); Organização Perceptiva (Completamento de Gravuras; Disposição de Gravuras, Cubos e Composição de Objectos) e Velocidade de Processamento (Código e Pesquisa de Símbolos). Todavia, os dados da investigação estão longe de uma convergência perfeita, havendo mesmo alguns autores, como é o caso de Kamphaus, 1993, citado por Watkins e Kuhs (2002), a referirem que os estudos de validade concorrente e preditiva, com amostras da população geral, não são convincentes quanto à existência inequívoca dos factores Resistência à Distracção e Velocidade de Processamento. Da mesma forma, as investigações que têm sido conduzidas no sentido de testar a validade preditiva destes dois factores isolados no desempenho académico (e.g., Glutting, Youngstrom, Ward, Ward e Hale, 1997) e nos distúrbios comportamentais (Oakland, Broom e Glutting, 2000) não têm recebido confirmação empírica. Por todos estes motivos, o valor dos índices Resistência à Distracção e Velocidade de Processamento devem ser interpretados com particular cautela.

No que se reporta à estrutura factorial das Escalas de Inteligência de Wechsler nos grupos com necessidades educativas especiais, embora sendo irrefutável a presença de alguns resultados inconsistentes, apontando para uma estrutura factorial específica nos grupos com necessidades educativas especiais (cf., por exemplo, Watkins e Canivez, 2001), a maior parte dos estudos, com as duas últimas versões da escala (WISC-R e WISC-III), apoiam a solução factorial encontrada na população geral (Kush, 1996; Konold, Kush e Canivez, 1997; Watkins e Kush, 2002). De referir que nesta investigação, o factor Resistência à Distracção apresenta uma fraca validade de constructo, confirmando resultados anteriores, como é o caso da investigação de Sullivan e Montoya, 1997, citada por Watkins e Kush (2002). A partir daí sugerem-se resultados característicos nestes índices ou factores nos sujeitos com dificuldades de aprendizagem, sendo que o perfil mais consistente aponta para pontuações progressivamente decrescentes ao longo dos índices Organização Perceptiva, Compreensão Verbal e Resistência à Distracção (cf. Daley e Nagle, 1996). Todavia, o dado ainda mais consensual reporta-se à presença de baixos resultados no 3º factor (Resistência à Distracção), havendo, no entanto, algumas reservas ou divergências quanto à natureza daquilo que mede esse factor. Assim, Reynolds e Kaufman (1990) referem que, a par da Resistência à Distracção, ele também poderá medir a memória a curto prazo, a aptidão simbólica e a ansiedade. Por seu turno, Wielkiewicz (1990, citado por Kush, 1996), acrescenta a este factor a aptidão numérica, enquanto Bannatyne (1974, citado por Kush, 1996) caracteriza este factor como medindo o processamento sequencial. De forma mais céptica, Watkins e Kuhs (2002), com

base num estudo de análise factorial confirmatória baseado nos desempenhos de 1201 alunos com D.A., sugerem que o factor Resistência à Distracção poderá ser um mero artefacto estatístico sem qualquer significado substantivo.

Os dados até agora revistos, são no entanto contrariados por algumas investigações ao revelarem uma estrutura factorial diferenciada nas amostras constituídas por crianças com dificuldades de aprendizagem. É o caso da investigação de Kush (1996), levada a cabo numa amostra de 327 estudantes com distúrbios de aprendizagem e pertencentes a diversas comunidades étnicas e culturais. Os resultados da análise factorial indicam a presença de três factores (o mesmo número e composição dos factores da amostra normativa portuguesa): Compreensão Verbal (Informação, Semelhanças, Aritmética, Vocabulário e Compreensão); Organização Perceptiva (Completamento de Gravuras, Disposição de Gravuras, Cubos e Composição de Objectos); e a Velocidade de Processamento (Código e Pesquisa de Símbolos). Os dois primeiros factores explicam 39% da variância dos resultados, valor que é muito próximo daquele que foi encontrado na amostra de aferição americana (43%). O terceiro factor apenas explica 5% da variância dos resultados e, muito provavelmente, não poderá assumir-se como um factor independente. Saliente-se que o factor Resistência à Distracção não emerge nesta solução factorial: o teste de Aritmética, que na amostra de aferição saturava no factor Resistência à Distracção é, nesta amostra, incluído no factor Compreensão Verbal; por sua vez, a Memória de Dígitos aparece como um teste isolado e com saturações muito baixas e idênticas nos três factores.

Mishra, Lord e Sabers (1989) também referem soluções factoriais diferenciadas nos grupos especiais da população, nomeadamente nas crianças com D.A. ou sobredotadas.

O estudo de Golombok e Rust (1992), conduzido no âmbito da aferição inglesa da WISC-III, contraria, de algum modo, os resultados das suposições mais comuns. Nele se verifica que o grupo de crianças com distúrbios de aprendizagem moderados ou severos, obtém a pontuação média mais elevada no factor Velocidade de Processamento (Código e Pesquisa de Símbolos), contrariando aparentemente a ideia de que os alunos com dificuldades de aprendizagem são mais distraídos ou têm dificuldades na focalização da atenção. Por seu turno, o grupo com dificuldades específicas na leitura apresenta resultados médios mais elevados nos factores Organização Perceptiva (104.7) e Compreensão Verbal (100.2), enquanto os valores mais baixos se registam nos factores Velocidade de Processamento (95.4) e Resistência à Distracção (93.2). Estes últimos, são resultados igualmente inesperados, nomeadamente os que respeitam ao índice Compreensão Verbal, se atendermos ao

facto de que as crianças com dificuldades específicas de leitura apresentam, com muita frequência, perturbações de ordem linguística (e.g., Reid, 1998; Newby, Recht e Caldwell, 1993).

Face a estes resultados, Kush (1996) conclui que, tal como na população normal, na população com dificuldades de aprendizagem, o QIEC (QI da Escala Completa), o QIV (QI Verbal) e o QIR (QI de Realização), podem ser interpretados como índices de inteligência relativamente independentes. As análises que se reportam aos restantes factores, tornam menos sustentáveis as interpretações que se possam fazer, pois a estrutura factorial da escala nesta população é substancialmente diferente daquela que se obtém na amostra de aferição.

Os índices de dispersão dos resultados

A análise da dispersão dos resultados nas escalas de Wechsler adquiriu uma relativa popularidade entre os psicólogos, sobretudo quando na década de setenta muitos estudos demonstraram a relação entre esses índices e certos distúrbios psicoeducativos (Clements, 1966; Dean, 1977; Waugh e Bush, 1977). Por exemplo, Rourke, 1998, citado por Kaufman e Lichtenberger (2000) julga ser possível definir subgrupos de dificuldades de aprendizagem combinando o valor da discrepância entre as subescalas verbal e realização com a informação complementar, recolhida por outros meios. Assim, as crianças com dificuldades de aprendizagem associadas a perturbações de ordem fonológica terão um QIV inferior ao QIR em pelo menos 10 pontos; por seu turno, as crianças com dificuldades de aprendizagem não-verbais invertem o padrão anterior: QIR inferior ao QIV, em pelo menos 10 pontos.

As metodologias de análise para calcular os índices de dispersão são diversificadas. Como mais frequentes apresentam-se: (i) a análise da discrepância entre o QIV e o QIR; (ii) a análise da dispersão dos resultados em torno da média individual (*scatter*), em que se compara o resultado obtido em cada subteste com o valor médio alcançado no conjunto dos testes (habitualmente utiliza-se como valor crítico o desvio-padrão do subteste – 3 pontos – ou apenas metade desse valor); (iii) a subtração do valor alcançado no subteste mais fraco ao valor atingido no subteste mais alto.

Progressivamente, os resultados da investigação têm levantado sérias reservas à eficácia destas análises e recomendam interpretações mais cautelosas (Anderson, Kaufman e Kaufman, 1976; Gutkin, 1979; Kaufman, 1981; citados por Patchett e Stansfield, 1992).

Assim, Kaufman (1981) depois de fazer uma revisão dos estudos sobre a discrepância entre as subescalas e a dispersão dos resultados na WISC-R em crianças com D.A., conclui que, apesar de se verificarem algumas diferenças com significado estatístico entre os grupos com e sem D.A., essas diferenças são tão reduzidas que deverão ter pouca validade em termos de diagnóstico diferencial.

Por seu turno, Kush (1996) numa amostra de 327 estudantes com D.A. não constata qualquer diferença entre subescalas (QIR= 93.5 *versus* QIV= 90.28). No mesmo sentido vão os resultados de Dumont e Willis (1995), ao observarem níveis de *scatter* semelhantes em dois grupos de crianças, respectivamente com e sem problemas de aprendizagem.

Neste panorama, as crianças sobredotadas com dificuldades de aprendizagem parecem constituir uma excepção, ao apresentarem sistematicamente índices de dispersão elevados (e.g., Kaufman e Kaufman, 1981; Schiff e cols., 1981, citados por Patchett e Stansfield, 1992). No entanto, é legítimo questionarmos se essa dispersão (frequentemente acompanhada por uma grande discrepância entre as subescalas verbal e de realização, em favor da primeira) deverá ser atribuída às dificuldades de aprendizagem ou se é um fenómeno estatístico. Com efeito, uma constatação frequente nos estudos efectuados com a WISC-R, é a de que o aumento dos índices de dispersão acompanha o crescimento registado no QI (Patchett e Stansfield, 1992). Por exemplo, a proporção de casos que apresentam diferenças significativas entre as subescalas Verbal e de Realização sofre uma progressão constante com o aumento do QI (Kaufman, 1976, citado por Patchett e Stansfield, 1992). Ou seja, face a estes resultados, é previsível assistir-se a um aumento dos índices de dispersão nas crianças sobredotadas com dificuldades de aprendizagem, sem que isso constitua um elemento importante para o diagnóstico diferencial.

No sentido de testar essa hipótese, Patchett e Stansfield (1992) levaram a cabo uma investigação, na qual analisaram a discrepância entre as subescalas e outros índices de dispersão dos resultados (e.g., subtracção do resultado mais baixo ao resultado mais elevado) num grupo de 290 crianças com 9 anos de idade e que apresentavam QIs a oscilarem entre 100 e 140. Tendo como referência o QI constituíram cinco grupos: grupo 1, QIs entre 100 e 109; grupo 2, QIs entre 110 e 119; grupo 3, QIs entre 120 e 129; grupo 4, QIs entre 130 e 139; grupo 5, QI > 140.

No que se refere à variabilidade entre subescalas não se registaram diferenças entre os cinco grupos. O mesmo não aconteceu nos outros índices de dispersão, em que se observaram diferenças inter-grupos significativas. Particularmente saliente, é o caso do índice de dispersão obtido através da sub-

tracção do resultado no teste mais baixo ao resultado no teste mais elevado. Nesta situação, os grupos 4 e 5 (com QI mais elevado) contrastam perfeitamente com os restantes, apresentando uma dispersão mais elevada. Em face destes resultados, os autores concluem que a dispersão dos resultados é um fenómeno estatístico que acompanha o aumento do QI e, sendo assim, o índice de dispersão perderá o valor clínico na identificação de determinados grupos especiais ou, pelo menos, deverá ser interpretado com mais reservas. A investigação conduzida por Pereira, Seabra-Santos e Simões (2003), numa amostra de crianças sobredotadas sem dificuldades de aprendizagem, reforça a interpretação de Patchett e Stansfield (1992).

O efeito Flynn: comparações WISC-R/WISC-III

Um outro dado que importa analisar, quando se utilizam as Escalas de Wechsler para avaliar a inteligência das crianças pertencentes a grupos especiais da população, refere-se ao efeito Flynn (Flynn, 1984, 1987). Este prevê um ganho médio de três pontos no QI dos sujeitos, em cada 10 anos que passam sobre a aferição de um teste de inteligência. No que se refere às Escalas de Inteligência de Wechsler a comparação dos resultados com a mesma versão do teste em dois momentos temporais distintos aponta para uma estabilidade moderada ou alta. Quando as medidas se efectuam com duas versões diferentes da escala (WISC-R e WISC-III) os resultados mostram descidas de 5 a 8 pontos nos valores de QI de Escala Completa (e.g., Bolen e Col., 1995; Graf e Hinton, 1994, citados por Truscott e Frank, 2001). Trata-se de um decréscimo esperado pelo efeito Flynn, já que aproximadamente duas décadas separam as duas versões americanas da escala (1974-1991). Os psicólogos devem estar atentos à intervenção deste efeito, pois ele tem implicações directas no diagnóstico das D.A.. Isto é, a inflação dos resultados nas escalas de QI acentua o valor da discrepância entre essa medida e os resultados nos testes aferidos de conhecimentos escolares, conduzindo a um aumento de falsos positivos. Ou seja, no momento em que passam a vigorar as normas da nova aferição, os resultados na escala de QI diminuem e, consequentemente, vai haver um número considerável de sujeitos que, até então eram considerados alunos com D.A., e que deixam de o ser a partir daquele momento, uma vez que já não respeitam o critério da discrepância. O conhecimento deste fenómeno é ainda mais importante no contexto norte-americano, sobretudo pelas implicações que acarreta a nível administrativo. Isto é, num sistema educativo em que os alunos beneficiários das medidas de educação especial têm de

“certificar” a permanência do seu *handicap* em cada triénio, submetendo-se a uma nova avaliação psicológica, pode acontecer que, num curto espaço de tempo (imediatamente após a nova aferição da escala), um determinado departamento estadual tenha que alterar substancialmente o quadro dos seus recursos logísticos e humanos.

Os mecanismos de acção do efeito Flynn ainda não são totalmente conhecidos, contudo a investigação tem demonstrado que ele é condicionado pela natureza dos instrumentos de avaliação e, também, que opera de modo diferenciado nos grupos especiais da população, afectando sobretudo aqueles que apresentam resultados afastados dos valores médios. Ou seja, as diferenças entre os QIs obtidos com a WISC-R e os QIs alcançados com a WISC-III serão mais restritos perto da distribuição central, aumentando à medida que nos aproximamos de condições mais extremas (deficiência mental e sobredotação).

No que respeita à natureza dos instrumentos de avaliação, tem-se verificado que a inflação prevista por Flynn afecta mais as pontuações dos testes não verbais, como é o caso das Matrizes de Raven, do que as pontuações dos testes predominantemente verbais (Truscott e Frank, 2001). Esta mesma tendência deveria registar-se nas Escalas de Inteligência de Wechsler, prevendo-se que nas sucessivas revisões haja descidas mais acentuadas na subescala de realização do que na subescala verbal. No entanto, isso nem sempre se verifica, pelo menos de modo assinalável. Assim, Weiss, 1995, citado por Carlton e Sapp (1997), numa análise de 22 estudos, constata pequenas variações, assistindo-se a um decréscimo médio de 3.2 pontos no QIV e de 4.4 pontos no QIR. Curioso, é o facto de que no conjunto dos três testes em que se registam as inflexões mais pronunciados (Semelhanças, Compreensão e Código), dois são parte integrante da subescala verbal. Por seu turno, Graf e Hinton, 1994 (citados por Truscott e Frank, 2001), observam que o efeito Flynn se circunscreve à subescala de realização.

Relativamente à acção diferenciada do efeito Flynn nos grupos especiais da população, os resultados também não são conclusivos.

Wechsler (1991) ao comparar as duas versões da escala num grupo de crianças com deficiência mental, observa que a evolução dos desempenhos da WISC-R para a WISC-III sofre quebras médias de 8.9 pontos no QIV, 6.8 pontos no QIR e 8.9 pontos no QIEC. Trata-se de um decréscimo mais acentuado do que o previsto pelo efeito Flynn, pois os 17 anos que separam as duas versões da escala faziam prever descidas médias de 5.7 pontos. O mesmo fenómeno é observado por Golombok e Rust (1992) em Inglaterra, numa amostra de 43 crianças com dificuldades de aprendizagem moderadas e se-

veras. De facto, a aplicação das duas versões das escalas de Wechsler em dois tempos mediados por um intervalo de 2 anos, regista um decréscimo da WISC-R para a WISC-III de 8.9 no QIV, 6.8 no QIR e 8.9 no QIEC.

Por seu turno, os resultados da investigação de Sapp, Abbot e Hinckley (1997), confirmam o efeito Flynn, mas não constata o seu carácter mais pronunciado nos grupos extremos. Assim, num grupo de alunos do ensino especial (com distúrbios de aprendizagem, deficiência mental, distúrbios de comportamento e outros problemas de saúde) registam um decréscimo médio de 5.36 pontos no QIEC, 5.18 pontos no QIV e 5.2 pontos no QIR. Saliente-se que neste estudo a inflexão é mais acentuada nos testes de Semelhanças, Vocabulário, Disposição de Gravuras e Cubos.

É importante referir que a nossa prática clínica tem demonstrado que o efeito Flynn se faz sentir de forma pronunciada na população com deficiência mental, mas não produz efeitos no grupo das crianças sobredotadas, que, paradoxalmente, invertem frequentemente a tendência desse fenómeno.

Graf e Hinton, citados por Truscott e Frank (2001), num estudo comparativo entre os resultados da WISC-R e da WISC-III, registam resultados que contrariam totalmente a norma esperada. Assim, verificam que o efeito Flynn se faz sentir no grupo de crianças com QI médio, mas não se regista no grupo de crianças com QI inferior a 90. Ou seja, na amostra de QI médio observam-se as descidas previstas pelo efeito Flynn enquanto no grupo de crianças que frequentam programas de educação especial o decréscimo médio é de apenas 2 pontos, ficando muito aquém dos 5.7 previstos. Destaque-se ainda, que no grupo com QI inferior há casos em que os resultados da WISC-III são superiores aos da WISC-R.

Negando também qualquer especificidade à actuação do efeito Flynn na população com necessidades educativas especiais, surgem os resultados da investigação de Weiss, 1995 (citado por Sapp, Abbott e Hinckley, 1997). Este autor analisou 22 estudos e constata que as diferenças no QIEC entre as duas escalas variam entre 1.35 e 18.0 pontos e que nos grupos especiais da população (dificuldades de aprendizagem, deficiência mental, etc.) se observa a mesma tendência.

No sentido de esclarecer melhor este fenómeno, Truscott e Frank (2001), estudaram o efeito Flynn numa amostra de 171 crianças com distúrbios de aprendizagem (predominantemente dificuldades específicas na aprendizagem da leitura), que num primeiro momento foram avaliadas com a WISC-R e, três anos depois, com a WISC-III. Os resultados mostram que os valores do QIEC alcançados nas duas versões da escala declinam da primeira para a segunda avaliação. Concluem que o efeito Flynn opera nas crianças com distúrbios de

aprendizagem, ainda que os valores observados sugiram que ele não produz os mesmos efeitos nas diferentes competências cognitivas, pois regista-se um decréscimo de 4.77 no QIEC, 2.93 no QIV e 6.68 no QIR.

Neste contexto vale a pena fazer referência a um estudo português que comparou os desempenhos na WISC e na WISC-III numa amostra de 70 crianças e adolescentes com a D. A. e/ou problemas de comportamento (Simões et al., 2003). Neste grupo clinicamente heterogéneo foram observadas diferenças importantes [28 pontos (QIEC), 18 pontos (QIV) e 31 pontos (QIR)] sistematicamente favoráveis à WISC. Note-se que existe uma diferença de aproximadamente 35 anos entre as datas de aferição da WISC (1968) e da WISC-III, e que a comparação habitual é entre a WISC-III e a WISC-R.

Face aos resultados, de um modo geral, podemos concluir, provisoriamente, que o efeito Flynn actua de modo indiferenciado, nos diferentes grupos da população e que a sua acção é mais pronunciada nalguns testes. No entanto, é um fenómeno que deve ser tomado em conta, particularmente no grupo das crianças com dificuldades de aprendizagem, pois condiciona decisões importantes na vida escolar destes alunos, nomeadamente o acesso à frequência de programas específicos de ensino-aprendizagem.

Metodologia

A WISC-III foi aplicada a 54 crianças com dificuldades de aprendizagem (57.5% rapazes e 42.5% raparigas). Na sua esmagadora maioria pertencem aos estratos sócio-económicos mais baixos. Tivemos como referência a classificação de profissões utilizada no censo de 1991, distinguindo nove categorias: 1- membros dos corpos legislativos, quadros dirigentes da função pública; directores e quadros dirigentes de empresas (2.3%); 2- profissões intelectuais e científicas; 3- profissões técnicas intermédias (2,3%); 4- empregados administrativos (16.7%); 5- pessoal dos serviços de protecção e segurança, dos serviços pessoais e domésticos e trabalhadores similares (50%); 6- trabalhadores da agricultura e da pesca; 7- trabalhadores da produção industrial e artesãos (11.9%); 8- operadores de instalações industriais e máquinas fixas, condutores e montadores; 9- trabalhadores não qualificados da agricultura, indústria, comércio e serviço (16.7%). A zona de residência é em 94.4% dos casos predominantemente urbana, não havendo nenhuma criança a residir numa zona classificada como predominantemente rural. Em resumo, são dados que confirmam a tendência constantemente verificada noutros estudos. Isto é, as crianças com dificuldades de aprendizagem são na sua maioria

rapazes que pertencem a estratos sociais mais desfavorecidos (e.g., Carlton e Sapp, 1997; Watkins e Glutting, 2000; Kush, 1996; Truscott e Frank, 2001; Golombok e Rust, 1992). No entanto, refira-se que na nossa amostra as raparigas estão representadas em maior número do que é habitual.

Na constituição do grupo com dificuldades de aprendizagem, foram considerados os seguintes critérios: (i) crianças que os professores sinalizavam como tendo D.A. ligeiras ou moderadas e que, por isso, beneficiavam de um regime educativo especial, nomeadamente apoio pedagógico acrescido e/ou adequação na organização de classes ou turmas, recorrendo a uma redução do número efectivo de alunos. Quer isto dizer, que o critério de diagnóstico das D.A. foi meramente administrativo e não clínico: são alunos que estão contabilizados nas estatísticas oficiais independentemente da existência de um diagnóstico clínico. Sendo uma prática muito discutível, é no entanto a mais usual nas nossas escolas, de acordo com as informações que conseguimos recolher junto da Direcção Regional de Educação. Excluímos os casos de D.A. severas de modo a reduzirmos a percentagem de falsos positivos, que muito provavelmente deveria ocorrer, nomeadamente pela inclusão de crianças com deficiência mental ligeira; (ii) as D.A. poderiam ser generalizadas ou específicas (e.g., dislexia); (iii) a idade da criança deveria oscilar entre 8 e 11 anos. Com a imposição de uma idade mínima, pretendemos um maior rigor na constituição da amostra, incluindo apenas aqueles casos em que as D.A. têm um carácter mais permanente (pelo menos dois anos de frequência escolar), distinguindo-as dos outros em que as dificuldades são transitórias e associadas ao início da escolaridade. Ao impormos um limite superior, foi nossa intenção não diluir os sujeitos por um leque etário muito amplo, que forçosamente introduziria um conjunto de variáveis confundentes, susceptíveis de dificultar a interpretação dos resultados; (iv) finalmente, a criança não deveria apresentar défices sensoriais, motores, problemas de comportamento e/ou perturbações psiquiátricas (nem poderia estar a ser medicada por razões de ordem psiquiátrica ou distúrbio comportamental).

O grupo de controlo é constituído por 54 crianças, seleccionadas a partir da amostra normativa da aferição portuguesa da WISC-III, sem dificuldades de aprendizagem. A metodologia seguida nesta selecção foi a dos "pares idênticos" (Drew e Hardman, 1985). Por conseguinte, as crianças deste grupo foram escolhidas uma a uma, em função das afinidades com um dos sujeitos do grupo "dificuldades de aprendizagem" no que respeita à idade, género, zona de residência (predominantemente urbana ou moderadamente urbana) e nível socio-económico.

Serão objecto de análise as seguintes variáveis: (i) distribuição dos resultados nos 13 testes que compõem a escala; (ii) análise dos resultados por três agrupamentos de testes, constituídos com base na solução factorial da escala para a população portuguesa: Índice Compreensão Verbal, Índice Organização Perceptiva e Índice Velocidade de Processamento; (iii) análise dos resultados em duas medidas de dispersão: valor diferencial entre o QIV e o QIR; distribuição dos resultados em torno da média individual (*scatter*).

Análise dos Resultados

Tendo como referência a definição de D.A., já abordada anteriormente, prevê-se a ausência de diferenças significativas entre as duas amostras ao nível do QIEC. Ou seja, as crianças com D.A. devem apresentar valores médios de QI idênticos aos de um grupo retirado aleatoriamente da população. No entanto, os resultados apresentados no Quadro 1 não confirmam essa previsão ($M=77.6$, $DP=12.5$ versus $M=97.1$, $DP=13.18$; $F=64.12$, $P=0.000$). Saliente-se ainda, que apenas 38% (21 crianças) apresentam um QIEC igual ou superior a 80 (valor proposto como limiar de QI para formular um diagnóstico clínico de D.A.). Esta ocorrência poderá encontrar explicação a dois níveis: (i) carácter datado das normas dos instrumentos psicométricos anteriormente utilizados na identificação das D. A.; (ii) fragilidade dos critérios de sinalização/identificação, na maior parte das vezes operado exclusivamente pelos professores. É ainda importante salientar que esses constrangimentos são inultrapassáveis na actual conjuntura do nosso sistema de educação especial, que funciona em condições de óbvia precariedade e no qual as decisões são tomadas, na maior parte das vezes, com base em falsas suposições ou “intuições” que carecem de fundamentação científica e de dados pedagógicos criteriosos. Quer isto dizer, que o diagnóstico das crianças com D.A. nas nossas escolas é muito discricionário. Entendemos que esta situação poderá ser vantajosa para o sistema, sobretudo quando a sua política educativa é orientada por indicadores económicos (custos), mas é decerto altamente penalizadora para os alunos, a quem se nega o respeito à diferença e o acesso a condições especiais de ensino-aprendizagem a que têm direito. Isto é, num sistema educativo que encara as D.A. como um parente pobre do ensino especial e para o qual não devem (ou não podem) ser canalizados grandes recursos, é mais rentável ter um grande número de alunos classificados no grupo das D.A., que, na sua óptica errada, não carecem de muitos recursos, sejam eles humanos ou materiais, do que ter alunos a engrossar outras categorias nosográficas, nomeadamente atrasos de

desenvolvimento, às quais, por pressão social, o Ministério reconhece mais necessidades e, consequentemente, orienta mais recursos.

Quadro 1 - *Distribuição dos resultados na WISC-III*

	DIF. APRENDIZAGEM				CONTROLO				
	Média	D. P.	Mx	Mm	Média	D. P.	Mx	Mm	P
QIV	80,20	12,37	110	60	98,50	13,00	129	69	0,000
QIR	82,10	13,30	106	53	98,00	13,49	128	65	0,000
QIEC	77,60	12,15	103	54	97,10	13,18	125	63	0,000
ICV	81,60	13,53	113	58	98,70	12,51	131	72	0,000
IOP	83,40	13,75	107	52	98,20	13,11	132	69	0,000
IVP	86,60	12,75	119	57	99,80	15,36	137	59	0,000
C.G.	7,80	3,18	13	1	9,50	3,48	15	1	0,008
INF.	6,90	2,43	13	1	10,00	2,38	15	4	0,000
CÓD.	7,70	2,41	13	3	10,00	2,96	18	3	0,000
SEM.	7,10	2,86	14	2	10,00	2,92	19	5	0,000
D.GRAV.	7,50	2,71	12	1	10,00	2,39	17	6	0,000
ARIT.	7,20	2,06	13	1	9,90	2,52	17	5	0,000
CUBOS	6,90	2,73	13	1	9,80	2,89	18	3	0,000
VOC.	7,40	2,53	14	2	9,80	2,66	17	4	0,000
C.OBJ.	8,00	3,17	14	1	10,00	2,91	17	4	0,000
COMP.	7,70	2,56	14	2	9,70	2,43	15	5	0,000
P.SIMB.	7,70	2,86	14	1	9,90	3,07	17	2	0,000
M.DIG.	7,70	2,46	14	3	10,01	3,11	19	3	0,000
LAB.	8,80	3,02	18	4	10,00	2,68	16	5	0,031

Repare-se ainda que os resultados observados na amostra de controlo também não atingem o valor médio esperado. Dado que se trata de amostras emparelhadas, este facto sugere o efeito bem conhecido da influência negativa das desvantagens socioculturais, associadas a menores competências linguísticas, no desempenho em testes de inteligência, aspecto presente em ambas as amostras.

Acautelando estas limitações, continuemos a analisar a distribuição dos resultados médios obtidos nas duas amostras. Em primeiro lugar, é notório o melhor desempenho do grupo de controlo em todos os testes e em ambas as subescalas. Todas as diferenças observadas apresentam valor estatístico com elevado nível de significância. Em ambos os grupos as diferenças entre o QIV e o QIR são mínimas e não apresentam qualquer significado estatístico. Trata-se de um resultado algo surpreendente, sobretudo no que concerne à amostra D.A., pois os testes da subescala Verbal estão mais directamente relacionados com a informação escolar, domínio em que as crianças com D.A. se encontrariam, por definição, mais deficitárias.

Relativamente à distribuição dos resultados pelos diferentes testes, salienta-se, em primeiro lugar, a fraca amplitude, com resultados médios a oscila-

rem entre um mínimo de 6.9 nos testes de Cubos e Informação e 8.8 nos Labirintos. Na amostra de controlo, a distribuição dos resultados é ainda mais homogénea, tal como pode ser observado na Figura 1. Em termos absolutos os piores desempenhos das crianças com D.A. observam-se nos testes de Informação, Cubos, Semelhanças e Aritmética. O que pode reflectir a limitação metodológica, anteriormente apontada, pois estes testes parecem ter em comum o factor g (raciocínio abstracto), reforçando a hipótese de que o fraco rendimento destes sujeitos (a que os professores chamam *dificuldades de aprendizagem*) possa ter origem em défices intelectuais ou outros problemas de natureza cognitiva. As melhores *performances* encontram-se nos testes de Labirintos e Composição de Objectos.

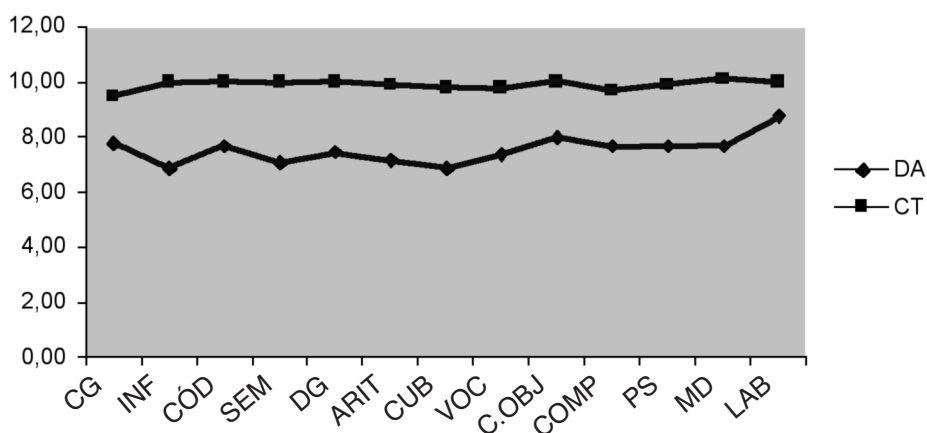


Figura 1 - Distribuição dos resultados na WISC-III

Na análise comparativa das duas amostras, saliente-se ainda que as diferenças mais acentuadas surgem nos testes de Informação, Semelhanças e Cubos (cf. Fig. 1). Mais uma vez, obtemos evidência que parece reforçar a hipótese de uma má sinalização das crianças com D.A., pois dois dos testes que melhor discriminam os grupos são saturados em factor g. Mais previsível, é a diferença observada no teste de Informação, dado que está relacionado com os conteúdos escolares.

Demandando a especificidade do perfil cognitivo das crianças com dificuldades de aprendizagem, vamos ensaiar uma segunda análise da distribuição dos resultados, recorrendo a um procedimento metodológico alternativo. Para o efeito, analisamos a distribuição dos resultados pelos três grupos de testes, constituídos com base na solução factorial encontrada na aferição portuguesa da WISC-III. Contudo, antes de avançarmos, é importante referir que

a solução factorial encontrada na população geral, poderá não se aplicar a grupos especiais, nomeadamente às crianças com D.A.. Lembramos que já anteriormente descrevemos estudos a apontarem para uma solução factorial diferenciada (Kush, 1996; Mishra, Lord e Sabers, 1989).

No Quadro 1, verificamos que a amostra das crianças com D.A. apresenta o valor mais elevado no conjunto dos testes que compõem o índice de Velocidade de Processamento (86.6, $dp = 12.75$); a segunda posição, é ocupada pelo grupo de testes que formam o índice de Organização Perceptiva (83.4, $dp = 13.75$); por fim, surgem os testes agregados no índice Compreensão Verbal (81.6, $dp = 13.53$). Refira-se que este perfil de desempenho contraria as previsões habitualmente apontadas para as crianças com D.A., definido por melhores desempenhos no índice de Organização Perceptiva (e.g., Daley e Nagle, 1996). Todavia o melhor desempenho no índice de Velocidade de Processamento, não constitui um dado isolado desta investigação, tendo já sido observado em estudos anteriores (por exemplo Golombok e Rust, 1992). A aplicação da análise da variância revela diferenças com significado estatístico entre as pontuações dos índices Velocidade de Processamento e Compreensão Verbal [$sig.=0.006$ ($t=2.848$; $gl=53$)]. Saliente-se ainda que a diferença entre o IVP e o ICV é de cinco pontos, ultrapassando a barreira dos três pontos, indicada por Wilkinson (1993) como critério discriminante do ponto de vista clínico.

Na amostra de controlo assistimos a uma hierarquização diferente, e menos pronunciada, dos três factores: $IVP > ICV > IOP$. A análise da variância não revela diferenças com significado estatístico.

Concluindo, os resultados observados conferem alguma validade a estas análises, como marcadores clínicos das dificuldades de aprendizagem.

Relativamente aos resultados aparentemente reduzidos do grupo com problemas de aprendizagem é necessário um outro tipo de comentários que é possível formular a partir da consulta do Quadro 2. Neste Quadro são apresentados os valores relativos aos intervalos de confiança (95%) dos resultados nos principais indicadores proporcionados pela WISC-III. Nele podemos constatar que os valores máximos do intervalo de confiança para os QI's oscilam entre 89 (QIV e QIEC) e 93 (QIR) no grupo com dificuldades de aprendizagem e entre 106 (QIV) e 107 (QIR) no grupo de controlo. Ou seja, quando consideramos os valores máximos do intervalo de confiança verificamos que o grupo com problemas de aprendizagem apresenta desempenhos muito próximos (QIV e QIEC) dos valores médios ou normais ou mesmo pontuações dentro deste parâmetro

(QIR). Os intervalos de confiança sinalizam o facto dos resultados envolverem sempre uma margem de erro pelo que é razoável considerarmos neste grupo os limiares superiores dos referidos intervalos, quer do ponto de vista da avaliação e do diagnóstico, quer da necessidade de intervenção.

Quadro 2 - *Intervalos de confiança (valores máximos e mínimos) relativos às pontuações médias dos QIs e Índices Factoriais*

	DIF. APRENDIZAGEM			CONTROLO		
	Mínimo	Máximo	Média	Mínimo	Máximo	Média
QIV	74	89	(80.2)	91	106	(98.5)
QIR	75	93	(82.1)	89	107	(98.0)
QIEC	71	89	(77.6)	89	106	(97.1)
ICV	75	91	(81.6)	91	107	(98.7)
IOP	76	95	(83.4)	89	108	(98.2)
IVP	79	100	(86.6)	89	111	(99.8)

No Quadro 3 apresentam-se os resultados referentes às duas medidas de dispersão.

Quadro 3 - *Distribuição dos resultados nas medidas de dispersão*

	DIF. APRENDIZAGEM				CONTROLO				
	Média	D. P.	Mx	Mm	Média	D. P.	Mx	Mm	P
SCATTER	1,80	0,38	2,85	1,09	1,90	0,37	2,63	1,11	0,289
QIV VS QIR	9,80	8,17	33,00	0,00	11,00	7,33	32,00	0,00	0,429

Na análise da dispersão tendo como medida o valor médio da dispersão intertestes (*scatter*), verificamos que não há diferenças significativas entre os dois grupos e que em ambas as amostras não se ultrapassa o desvio das duas unidades proposto por Wechsler como critério clínico para discriminar um desempenho normativo de um desempenho atípico (cf. Bourges, 1979; Marques, 1969). O facto de se observar uma dispersão ligeiramente mais elevada na amostra de controlo, relativiza ainda mais o valor clínico desta medida. É de salientar que este resultado, embora contrarie um elevado número de investigações, está longe de constituir um facto singular, pois a mesma tendência já foi observada num leque alargado de investigações anteriores (Kaufman, 1981; Kush, 1996; Patchett e Stansfield, 1992; Dumont e Willis, 1995).

Relativamente à diferença entre subescalas (*Verbal versus Realização*), constatamos valores normativos muito semelhantes nas duas amostras, o que contraria a ideia habitual de que estas crianças apresentam uma discrepância acentuada entre as duas subescalas. Refira-se que os resultados de investigações anteriores levadas a cabo noutros países também vão no mesmo sentido.

Assim, Carlton e Sapp (1997) numa amostra de alunos que frequentavam o ensino especial, e dos quais 83% apresentavam D.A., não observaram diferenças significativas entre as escalas Verbal e de Realização. Slate, 1995, citado por Kaufman e Lichtenberger (2000) ao examinarem a discrepância entre subescalas em grupos de sujeitos com deficiência mental, dificuldades específicas de aprendizagem e sem qualquer perturbação, observam que a discrepância a favor da subescala de Realização, está presente nas três amostras. Golombok e Rust (1992) nos estudos da aferição inglesa da WISC-III em amostras D.A. também verificam que os resultados da escala Verbal são inferiores aos da escala de Realização, mas sem discrepâncias assinaláveis.

Discussão dos resultados e limitações da presente investigação

Os resultados demonstram que as crianças com dificuldades de aprendizagem apresentam um perfil que se caracteriza sobretudo pela fraca variabilidade dos desempenhos nos treze testes e nas duas subescalas que constituem a WISC-III.

Em termos absolutos os testes mais fracassados são a Informação, Semelhanças e Cubos, sugerindo algum grau de comprometimento intelectual dos sujeitos da nossa amostra.

O agrupamento de testes, de acordo com a solução factorial encontrada na amostra de aferição, revela o protagonismo do índice de Velocidade de Processamento, o que, de algum modo, parece denunciar a razoável capacidade das crianças com D.A. mobilizarem a sua atenção. Tal facto contrasta com os resultados de um grande número de investigações, que apontam para problemas na atenção dirigida. Como explicar este paradoxo? Julgamos que os testes agrupados no factor Velocidade de Processamento (Código e Pesquisa de Símbolos) ainda que mobilizem a capacidade de atenção, distinguem-se pelo facto de solicitarem um envolvimento activo (motor) do sujeito e requererem o processamento visual da informação, capitalizando a eficácia do seu funcionamento cognitivo, que é deficitário sobretudo no processamento auditivo da informação, ao impor-lhes uma atitude mais passiva. Neste contexto, poderemos especular que as crianças com dificuldades de aprendizagem apresentam um estilo cognitivo diferenciado e que só capitalizam os seus recursos quando as tarefas que lhes são propostas respeitam essa especificidade. Dito de outro modo, uma criança com D.A. poderá melhorar o seu processo de aprendizagem quando os estímulos lhe são apresentados por via visual,

pois esta facilita-lhe a captação mais exacta da informação e a sua recolha organizada.

Por último, a hierarquização diferente e acentuada dos índices factoriais na amostra das crianças com D.A., parece constituir prova do valor clínico das escalas de Wechsler como marcador das dificuldades de aprendizagem, conferindo-lhe algum valor na previsão de distúrbios psicopedagógicos ou para operar ao nível do diagnóstico diferencial.

Por último, importa questionar o significado dos resultados obtidos, clarificando-os à luz das limitações que a investigação apresenta. Como já referimos anteriormente, o critério que presidiu à constituição da amostra de crianças com dificuldades de aprendizagem, revelou-se muito frágil, em virtude dos problemas que o nosso sistema de educação especial apresenta. Em próximas investigações a amostragem deverá ser constituída exclusivamente com base em critérios clínicos e diferenciar os diversos subtipos de D.A.. De todo o modo, parece-nos importante sensibilizar os psicólogos para a necessidade destes trabalharem, sistematicamente, com os intervalos de confiança dos resultados, valorizando os limiares superiores para efeitos de análise e interpretação.

Referências

- Carlton, M. e Sapp, G. (1997). Comparison of WISC-R and WISC-III scores of urban exceptional students. *Psychological Reports*, 80, 755-760.
- Daley, C. e Nagle, R. J. (1996). Relevance of the WISC-III indicators for assessment of learning disabilities. *Journal of Psychoeducational Assessment*, 14, 320-333.
- Dumont, R. e Willis, J. O. (1995). Intrasubtest scatter on the WISC-III for various clinical samples vs. the standardization sample: An examination of WISC folklore. *Journal of Psychoeducational Assessment*, 13, 271-285.
- Flanagan, D. e Alfonso, V. (1993). WIAT subtest and composit predicted-achievement values based on WISC-III verbal and performance IQs. *Psychology in the Schools*, 30, 310-320.
- Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, 101, 171-191.
- Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, 95, 29-51.
- Gersons-Wolfensberger, D. e Ruijsenaars, J. (1997). Definition and treatment of dyslexia: A report by the Committee on Dyslexia of the Health Council of Netherlands. *Journal of Learning Disabilities*, 30 (2), 209-213.
- Glutting, J. J., Youngstrom, E. A, Ward, T., Ward, S. e Hale, R. L. (1997). Incremental efficacy of WISC-III factor scores in predicting achievement: What they tell us? *Journal of Psychological Assessment*, 9, 295-301.
- Golombok, S. e Rust, J. (1992). *Wechsler Intelligence Scale for Children—Third Edition UK*. London: The Psychological Corporation.
- Kamphaus, R. W. (1993). *Clinical assessment of children's intelligence*. Boston: Allyn and Bacon.
- Kaufman, A. e Lichtenberger, E. (2000). *Essentials of WISC-III and WPPSI-R assessment*. New York: John Wiley e Sons, Inc.
- Kaufman, A. S. e Lichtenberger, E. (1999). *Essentials of WAIS-III assessment*. New York: John Wiley e Sons, Inc.
- Kaufman, A. S. (1994). *Intelligent testing with the WISC-III*. New York: John Wiley e Sons, Inc.
- Kaufman, A.S. (1979). *Intelligent testing with the WISC-R*. New York: John Wiley e Sons, Inc.
- Kaufman, A. S. (1975). Factor analysis of the WISC-R at eleven age levels between 6 and 16 years. *Journal of Clinical and Consulting Psychology*, 43, 135-147.
- Konold, T. R., Kush, J. C. e Canivez, G.L. (1997). Factor replication of the WISC-III in three independent samples of children receiving special education. *Journal of Psychoeducational Assessment*, 15, 123-137.

- Kush, J. (1996). Factor structure of the WISC-III for students with learning disabilities. *Journal of Psychoeducational Assessment*, 14, 32-40.
- Lima, P., Dores, A. P. e Costa, A. F. (1991). Classificações de profissões nos censos 91. *Sociologia: Problemas e Práticas*, 10, 43-66.
- Lustberg, R., Motta, R. e Naccari, N. (1990). A model using the WISC-R to predict success in programs for gifted students. *Psychology in the Schools*, 27, 126-131.
- Misha S., Lord, J. e Sabers, A. (1989). Cognitive processes underlying WISC-R performance of gifted and learning disabled navajos. *Psychology in Schools*, 26, 31-36.
- Newby, R., Recht, D. e Caldwell, J. (1993). Validation of a clinical method for the diagnosis of two subtypes of dyslexia. *Journal of Psychoeducational Assessment*, 11, 72-83.
- Oakland, T., Broom, J. e Glutting, J. (2000) Use of freedom from distractibility and processing speed to assess children's test-taking behaviors. *Journal of School Psychology*, 38, 469-475.
- Patchett, R. e Stansfield, M. (1992). Subtest scatter on the WISC-R with children of superior intelligence. *Psychology in the Schools*, 29, 5-11.
- Phelps, L. (1989). Comparison of scores for intellectually gifted students on the WISC-R and the Fourth edition of the Stanford-Binet. *Psychology in Schools*, 26, 125-129.
- Prifitera, A. e Dersh, J. (1993). Base rates of WISC-III diagnostic subtest patterns among normal, learning disabled, and ADHD samples. *Journal of Psychoeducational Assessment*, WISC-III Monograph, 43-55.
- Prifitera, A., Weiss, L. G. e Saklofske, D. (1998). *WISC-III clinical use and interpretation: Scientist-practitioner perspectives*. New York: Academic Press.
- Reid, G. (1998). *Dyslexia: A practitioner's handbook*. London: John Wiley e Sons.
- Reynolds, C. R. e Kaufman, A. S. (1990). Assessment of children's intelligence with Wechsler Intelligence Scale for Children - Revised (WISC-R). In C. R. Reynolds e R. W. Kamphaus (Eds.), *Handbook of Psychological and educational assessment of children: Intelligence and achievement* (pp. 127-165). New York: Guilford.
- Robinson, E. e Nagle, R. (1992). The comparability of the test of cognitive skills with the Wechsler Intelligence Scale for Children-revised and the Stanford-Binet: Fourth Edition with gifted children. *Psychology in Schools*, 29, 107-112.
- Rourke, B. P. (Ed.) (1991). *Neuropsychological validation of learning disability subtypes*. New York: Guilford Press.
- Saccuzzo, D. P., Johnson, N. e Russell, G. (1992). Verbal versus performance IQs for gifted african-american, caucasian, filipino, and hispanic children. *Psychological Assessment*, 4 (2), 239-244.
- Sapp, G., Abbot, G. e Hinckley, R. (1997). Examination of the validity of the WISC-III with urban exceptional students. *Psychological Reports*, 81, 1163-1168.

- Siegel, L. S. (1989). IQ is irrelevant to the definition of learning disabilities. *Journal of Learning Disabilities*, 22, 469-479.
- Simões, M. R., Azevedo, A., Marques, B., Morgado, C., Costa, C., Marques, L. e Lopes, A. F. (2004). Relações entre os desempenhos nas versões portuguesas da WISC e da WISC-III num grupo de crianças e adolescentes com dificuldades de aprendizagem e/ou problemas de comportamento. *Psicologia e Educação*, 2(1), 19-40.
- Truscott, S. D. e Frank, A. J. (2001). Does the Flynn effect affect IQ scores of students classified as LD? *Journal of School Psychology*, 39 (4), 319-334.
- Ward, S. B.; Ward, T. J.; Hatt, C.; Young, D. e Mollner, N. (1995). The incidence and utility of the ACID, ACIDs and SCAD profiles in a referred population. *Psychology in the Schools*, 32, 276-276.
- Watkins, M. e Canivez, G. (2001). Longitudinal factor structure of the WISC-III among students with disabilities. *Psychology in the Schools*, 38 (4), 291-299.
- Watkins, M. e Glutting, J. (2000). Incremental validity of WISC-III profile elevation, scatter, and shape information for predicting reading and math achievement. *Psychological Assessment*, 12 (4), 402-408.
- Watkins, M., Kush, J. e Glutting, J. (1997). Discriminant and predictive validity of the WISC-III among children with learning disabilities. *Psychology in the Schools*, 34 (4), 309-319.
- Watkins, M. e Kush, J. (2002). Confirmatory factor analysis of the WISC-III for students with learning disabilities. *Journal of Psychoeducational Assessment*, 20, 4-19.
- Wechsler, D. (2003). *Escala de Inteligência de Wechsler para Crianças – Terceira Edição*. Lisboa: Cegoc.
- Wechsler, D. (1991). *Manual for the Wechsler Intelligence Scale for Children—Third edition*. New York: The Psychological Corporation.
- Wechsler, D. (1996). *Échelle d'Intelligence de Wechsler pour Enfants: Troisième édition. Manual*. Paris: Les Editions du Centre de Psychologie Appliquée.
- Weiss, L. G. e Prifitera, A. (1995). An evaluation of differential prediction of WIAT achievement scores from WISC-III FSIQ across ethnic and gender groups. *Journal of School Psychologist*, 33 (4), 297-303.
- Wilkinson, S. C. (1993). WISC-R profiles of children with superior intellectual ability. *Gifted Child Quarterly*, 37 (2), 84-91.

INTELLIGENCE ASSESSMENT IN LEARNING DISABILITIES: RESEARCH WITH WISC-III

Marcelino Pereira

Mário R. Simões

Serviço de Avaliação Psicológica, Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra

Abstract: This paper aims to establish a reference for interpreting scores of learning disabilities children on the WISC-III (Wechsler Intelligence Scale for Children - Third Edition). In order to achieve this goal, a sample of 54 children whose teachers screening as presenting learning disabilities was compared to a control group, using the matched pairs methodology. The following variables are analyzed: (i) distribution of scores on the thirteen subtests that compose the scale; (ii) magnitude and comparison of scores along the three factors (Verbal Comprehension, Perceptual Organization and Processing Speed); (iii) results on the following two scatter measures: discrepancy between verbal IQ and Performance IQ; distribution of subtest scores around the individual mean.

KEY-WORDS: *WISC-III, Intelligence Assessment, Learning Disabilities.*

INTELIGÊNCIA, EDUCAÇÃO E SUCESSO: AS ABORDAGENS DIFERENCIAL E SISTÊMICA DA INTELIGÊNCIA HUMANA E SUAS IMPLICAÇÕES PARA A EDUCAÇÃO

Maria João Afonso

Faculdade de Psicologia e de Ciências da Educação, Universidade de Lisboa, Portugal

Resumo

Este artigo aborda a noção de sucesso e a sua relação com a inteligência no âmbito de dois paradigmas de investigação e avaliação: o paradigma diferencial, subjacente às técnicas tradicionais de medida da inteligência, e o paradigma sistémico, aqui representado por uma das mais importantes teorias da actualidade, a Teoria Triárquica da Inteligência Humana de Robert Sternberg. São identificadas algumas implicações da teoria, e do conceito de Inteligência Funcional (Successful Intelligence) a que deu origem, para a educação e a avaliação da inteligência. O Teste Triárquico de Aptidões de Sternberg – Revisto é apresentado como operacionalização da teoria e são discutidos alguns resultados obtidos em Portugal com a versão experimental portuguesa (nível H – ensino superior e adultos) (Afonso, 2004): consistência interna e análise factorial confirmatória.

PALAVRAS-CHAVE: *Inteligência, inteligência funcional, STAT-R (H), análise factorial confirmatória.*

Charles Darwin e Albert Einstein eram homens inteligentes? A pergunta soa absurda, de tal modo a resposta é óbvia e imediata. Enquanto cientistas, apenas nos ocorre considerá-los “homens de sucesso”, pelo lugar destacado que conquistaram nos respectivos domínios, pelo legado de saber que deixaram à Humanidade, pela forma indelével como mudaram os nossos conceitos de Homem e do seu lugar no Universo. Assim se converteram em dois casos de grande imortalidade, na acepção de Kundera. E, contudo, nem um nem outro deixou de viver experiências de insucesso. Nas palavras do próprio Charles Darwin (1809-1882) é possível discernir essa vivência, em ligação com a sua escolarização:

“A escola como meio de educação, para mim foi um simples vazio. Durante toda a minha vida, sempre fui simplesmente inca-

paz de lidar eficazmente com qualquer matéria linguística. [Na escola] era dada especial importância à construção de versos, algo que eu jamais consegui fazer bem (...). Quando deixei a escola (...) estou convencido de que era considerado, por todos os meus professores e pelo meu próprio pai, um rapaz muito vulgar, mesmo um pouco inferior ao nível intelectual normal. (...) as únicas qualidades que neste período pareciam promissoras eram os meus interesses diversificados e vinculados, o grande zelo com que me dedicava a qualquer tema que me interessasse e o genuíno prazer que sentia na compreensão de qualquer assunto ou fenómeno complexos." (Darwin, 1905, pp. 29-30)

Apesar da crença, mais ou menos generalizada mas infundada, no insucesso escolar de Einstein (1879-1955), não foi ao longo da escolaridade básica, diferentemente de Darwin, que enfrentou o insucesso, sendo que era até, de acordo com Abraham Pais (1982/2004), "um aluno acertadinho, persistente, mas lento, que resolvia os problemas de matemática com segurança, mas não sem erros de cálculo" (pp.57-58). Veio a tornar-se "o melhor aluno na escola" (p.59), de quem a própria mãe disse, em carta dirigida à avó de Einstein: "O Albert recebeu as notas. Foi novamente o melhor, a informação é brilhante" (p.58). Mesmo assim, se atentarmos ao supra referido relato biográfico, é assinalável a referência à apreensão da família, durante os primeiros anos de vida, relativamente a que aquela criança pudesse ser "atrasada", devido à configuração invulgar da cabeça e ao excessivo tempo que demorou a começar a falar, e pode supor-se também algum desajustamento à escola – "no conjunto, não gostou dos anos de escolaridade: professores autoritários, estudantes servis, ensino livresco, nada disso lhe caía bem. (...) Sentia-se isolado e fazia poucos amigos na escola" (p.58). É, no entanto, mais tarde, durante a adolescência e o início da idade adulta, que Einstein enfrenta alguns insucessos que o desencorajam: em 1895 reprova no exame de acesso ao curso de Engenharia Eléctrica no *Eidgenössische Technische Hochschule (ETH)* de Zurique, sentindo-se forçado a obter o diploma do ensino médio (*matura*) que veio a proporcionar-lhe esse acesso no ano seguinte, e em 1900, após a sua graduação, falhou a obtenção de um lugar como assistente no *ETH*, ao contrário de todos os outros estudantes graduados na mesma altura. A situação de desemprego prolongou-se e, apesar das múltiplas tentativas de candidatura a lugares universitários de assistente, em 1901 concluiu: "Abandonei a ambição de ingressar numa universidade..." (p.67). Só em Junho de 1902 assume finalmente as suas primeiras funções profissionais, a que acedeu com o

apoio de um amigo: perito técnico de terceira classe na repartição de patentes, em Berna. Funções bem diversas das que, nas suas próprias palavras, ambicionava aos 16 anos:

“(...) irei para o *ETH*, em Zurique. Ficarei lá quatro anos, para estudar matemática e física. Penso tornar-me professor naqueles ramos das ciências naturais, escolhendo a sua parte teórica. Eis as razões que me levaram a este projecto. Acima de tudo, a [minha] predisposição para o pensamento matemático abstracto, [e a minha] falta de imaginação e de capacidade prática.” (Einstein, *Mes projects d’avenir*, 1895 citado em Pais, 1982/2004, pp.61 e 64-65)

A “falta de capacidade prática”, de que mostrava consciência, terá sido em parte responsável pela sua dificuldade em conseguir o emprego a que aspirava? Possivelmente. E poderemos então concluir que Albert Einstein era um homem inteligente? Em rigor, a resposta a esta pergunta depende da forma como definimos “inteligência” e da relação que assumimos entre “inteligência” e “sucesso”.

Inteligência e Sucesso

Admite-se que a origem da palavra “inteligência” remonte ao período romano, sobretudo aos escritos de Cícero (106-43 a.C.): terá resultado da aglutinação de duas palavras, *inter* (entre) e *legere* (relacionar, estabelecer relações, escolher) (Miranda, 2000), e da derivada palavra latina *intelligentia*, que significava “entendimento, conhecimento” (Dicionário Houaiss da Língua Portuguesa, 2004). Enquanto objecto de reflexão filosófica, o conceito manteve ao longo da história da humanidade uma estreita ligação com os contextos sociais, políticos e ideológicos, assumindo a noção de inteligência em geral uma natureza consistente com os valores dominantes e com as qualidades humanas enaltecidas pela sociedade em cada época (ver por exemplo Richardson, 1991).

O estudo científico da inteligência data apenas dos finais do século XIX e emergiu no quadro de uma nova área das ciências naturais, fortemente influenciada pela Biologia, em particular pelas ideias evolucionistas de Charles Darwin: esta área, que veio a ser designada “psicologia diferencial”, admitia então o papel da inteligência no processo evolutivo da espécie humana, pelo que o conceito de inteligência se ligou desde logo, e de forma muito explícita,

à ideia de adaptação, de sucesso adaptativo, ideia nuclear na *Teoria da Evolução das Espécies*.

A “inteligência enquanto adaptação” tornou-se um conceito chave que de forma mais ou menos explícita atravessou a história da psicologia da inteligência humana, na base da investigação fundamental como da aplicada, no cerne do desenvolvimento teórico como das técnicas de avaliação da inteligência (Miranda, 2002, 2004). Este conceito emergiu em ambos os simpósios escritos sobre a natureza da inteligência realizados durante o século XX (Thorndike et al., 1921.; Sternberg e Detterman, 1986) e integrou as mais divulgadas e aceites concepções e definições de inteligência: constitui um dos três critérios do pensamento inteligente propostos por Binet e Simon em 1909 (Carroll, 1982, p. 36); reaparece na clássica definição de David Wechsler (1944, p.3) “a inteligência é o agregado ou capacidade global do indivíduo para agir finalizadamente, pensar racionalmente e proceder com eficiência em relação ao meio”; está subjacente ao paradigma construtivista da inteligência humana (adaptação é assimilação e acomodação) (Miranda, 2002, 2004); é retomado na proposta de uma definição sintética de inteligência por Sternberg (1982, p.3) “[*intelligence is*] *goal-directed adaptive behavior*” [comportamento adaptativo e direccionado].

Os primeiros testes de inteligência surgiram no virar do século XIX para o XX, num contexto de acentuada mudança nos planos social, económico, educacional e intelectual. A explosão demográfica, as implicações remotas da revolução industrial – a emergência de uma sociedade industrializada e, consequentemente, de uma nova organização do trabalho – o desenvolvimento da educação de massas e a democratização do ensino (Carroll, 1982) foram factores que contribuíram para estimular a criação de métodos destinados a identificar as capacidades dos indivíduos e prever a sua adaptação aos contextos escolar, profissional e, mais tarde, com o eclodir da I Grande Guerra, também ao contexto militar. Adaptação significava à época “rentabilidade” e as variáveis cognitivas, como a inteligência, eram consideradas os melhores preditores do “sucesso”, escolar como profissional. As sociedades ocidentais valorizavam então a “organização científica do trabalho” e a “colocação do homem certo no lugar certo” para aumentar a “prosperidade”, tanto para o empregador como para o empregado (Taylor, 1911); as diferenças individuais eram por isso vistas, à luz de um emergente *Darwinismo Social* (Hofstadter, 1944 citado em Carroll, 1982), como presumivelmente hereditárias e determinantes na adaptação dos indivíduos às exigências da nova sociedade industrializada e tecnológica (Afonso, 2005a).

O “sucesso”, objectivamente avaliável através das classificações escolares ou das avaliações de desempenho profissional, não só era significativamente correlacionado com o nível de inteligência – numerosas investigações nos meios escolar e laboral assim o documentavam, justificando a implantação de práticas generalizadas de avaliação cognitiva sempre que se pretendia a predição do sucesso – como constituía mesmo critério para a estimação empírica da validade das próprias técnicas de avaliação da inteligência (“validade relativa a um critério”: “concorrente” e “preditiva”). Daí o carácter mais fortemente empírico do que conceptual dos fundamentos da larga maioria das técnicas diferenciais de avaliação da cognição humana oriundas da primeira metade do século XX (Afonso, 1997, 2002).

A inteligência foi, assim, a primeira variável cognitiva a merecer a atenção dos psicólogos diferencialistas, como objecto de investigação e de medição psicológica, e as diferenças individuais fundamentaram o mais antigo paradigma, e a mais clássica metáfora, da investigação e da avaliação da inteligência humana: o paradigma diferencial (Miranda, 1986, 2000, 2002) e a metáfora geográfica (Sternberg, 1990). No “mapa mental” que consubstancia a metáfora geográfica, cuja unidade fundamental é o factor (Miranda, 2002), assume especial relevo uma aptidão ampla, abrangente de um largo espectro do funcionamento cognitivo, designada inteligência geral, ou *g*. Esta noção de *g* é por muitos considerada uma das grandes conquistas da psicologia científica (ver por exemplo Nyborg, 2003), pela quantidade e qualidade das evidências empíricas que a suportam e pelas implicações, no plano teórico, como no das aplicações. Ela atravessou, de facto, um século – de 1904, quando Charles Spearman pela primeira vez a introduziu até à actualidade – assume posição de destaque nos mais consensuais modelos estruturais das aptidões humanas (hoje sintetizados no chamado “modelo C-H-C” ou “Cattell-Horn-Carroll”) e está, de forma mais ou menos explícita, subjacente aos testes de “inteligência geral” ainda em plena utilização, e à sua aplicação em domínios diversificados de intervenção psicológica, entre os quais a educação.

Inteligência e Educação

Teve enorme importância histórica a primeira tentativa de aplicação de “testes mentais” ao domínio da educação: aconteceu em 1896 quando James McKeen Cattell postulou que estes testes, inspirados nas metodologias que conheceu e adoptara quando trabalhou com Galton, poderiam ser pre-

ditivos do sucesso académico dos estudantes. A verificação empírica, contudo, não sendo favorável a esta hipótese (Carroll, 1982) deu oportunidade a que o psicólogo francês Alfred Binet sugerisse que a medição da inteligência, como dimensão complexa que é, deveria basear-se na observação do indivíduo exposto a tarefas complexas, envolvendo actividade mental semelhante à quotidiana. Esta proposta constituiu um verdadeiro salto epistemológico, um salto significativo no nível de observação e análise dos fenómenos cognitivos (na acepção de Reuchlin, 1999/2002) que determinou o aparecimento, em 1905, daquele que é verdadeiramente considerado o primeiro “teste de inteligência”: a Escala Métrica de Inteligência para Crianças de Binet e Simon.

Esta escala, que nasceu com o propósito explicitamente ligado à educação de fazer o despiste das crianças “retardadas” à entrada para a escolaridade básica, integrava tarefas que procuravam permitir a predição da adaptação à escola, tarefas representativas do desempenho típico das crianças de determinada idade e que envolviam sobretudo a compreensão da linguagem e o raciocínio aplicado a conteúdos verbais, numéricos e figurativos. Pela ampla expansão internacional desta escala e das suas sucessivas edições (1908 e 1911), e pelas suas implicações ao nível da criação de novos testes baseados numa concepção de medida nelas inspirada, não é de surpreender que os testes de inteligência tenham vindo a assumir um carácter fortemente académico, sendo mesmo por vezes utilizados como testes de “aptidão escolar” (*scholastic aptitude*) (Carroll, 1982; Sternberg, 2003a).

Nalguns países, muito em particular nos EUA, generalizou-se ao longo do século XX a utilização de testes de “inteligência” e de “aptidão escolar” como preditores do sucesso académico e profissional – as correlações dos resultados em testes de inteligência (por exemplo QIs) com as classificações escolares, na ordem de .50, com o número de anos de escolaridade, na ordem de .55 e com o sucesso profissional, tipicamente entre .30 e .50 (Neisser et al., 1996), por muito tempo justificaram e fundamentaram tal utilização. Esta predição, necessária do ponto de vista da “rentabilização” de recursos humanos e materiais, tornou-se central em orientação escolar e profissional (Afonso, 2005a) e em selecção para a formação e o emprego. Compreende-se, assim, que nos EUA tenha surgido a partir da década de 50 um verdadeiro movimento de contestação aos testes (Carroll, 1982; Resnick e Resnick, 1982), incidindo especialmente nos testes de inteligência pelo impacto que tinham na vida dos cidadãos, ao limitarem com frequência as oportunidades de desenvolvimento, sobretudo em alguns grupos populacionais minoritários (*impacto adverso*); e compreende-se também que

tenha sido nesse mesmo país que, a partir da década de 70, surgiram no meio científico as posições mais fortemente críticas em relação à chamada “medida tradicional da inteligência”, a par de novas propostas de conceitualização e de medição da inteligência: Gardner (1983, 1999) e Sternberg (1985, 1996).

Os trabalhos de Robert Sternberg (Universidade de Yale) assumem uma particular importância no âmbito da educação, por proporem um modelo que procura ultrapassar o tradicional ciclo fechado entre os testes de inteligência – que avaliam as competências exigidas pela escola – e a escola – que promove o desenvolvimento das competências medidas pelos testes. De acordo com o autor, este ciclo fechado, que artificialmente favorece a emergência de correlações positivas entre os resultados dos testes de inteligência e o sucesso escolar, limita o âmbito das competências abrangidas pelo conceito tradicional de inteligência – apenas sendo contempladas as que se correlacionam com o sucesso acadêmico, como o pensamento analítico e a memória – e negligencia outras, imprescindíveis na adaptação ao “mundo real” (ao “sucesso”, também fora da escola) – como a capacidade para lidar eficazmente com situações novas, com tarefas de ordem prática, com as relações interpessoais ou com a gestão dos recursos pessoais. Ao promover um conjunto limitado de competências, quer a escola quer a avaliação psicológica estarão a condenar ao insucesso estudantes que, numa perspectiva mais ampla e complexa da inteligência, poderiam alcançar níveis adequados, ou até elevados, de adaptação e sucesso (Sternberg, 1998a, 2000, 2002b). Deste ponto de vista, ainda que os testes clássicos possam ser preditores do sucesso acadêmico, dificilmente poderão sê-lo em relação ao sucesso na vida quotidiana, já que este parece depender de um leque mais alargado e complexo de competências que excedem as que são promovidas pela escola. Sternberg propõe, assim, uma nova teoria – Teoria Triárquica da Inteligência Humana – e um novo conceito – Inteligência Funcional (*Successful Intelligence*), teoria e conceito com importantes implicações e aplicações na área da educação.

Teoria Triárquica da Inteligência Humana

A Teoria Triárquica da Inteligência foi formalmente apresentada por Robert Sternberg numa obra de título sugestivo: *Beyond IQ* (1985a). Nesta obra, que condensa muita da investigação e reflexão do autor desde os anos 70, aparece o primeiro esboço de uma teoria que tem vindo a ser empiricamente testada e sucessivamente aperfeiçoada. A sua designação remete para uma

organização em três subteorias que estabelecem a ligação entre a inteligência e 1) o mundo interno do indivíduo (*Subteoria Componencial*), 2) a experiência, adquirida ao longo do processo temporal de interacção entre os mundos interno e externo do indivíduo (*Subteoria Experiencial*) e 3) o mundo externo do indivíduo (*Subteoria Contextual*).

A *Subteoria Componencial* identifica os mecanismos mentais ou as componentes de processamento de informação subjacentes ao pensamento inteligente: as *metacomponentes* (processos executivos de ordem superior responsáveis pelo planeamento, pela monitorização e pela avaliação do funcionamento das componentes de nível inferior); as *componentes de execução* (processos de ordem inferior encarregues da resolução de problemas, de acordo com o planeamento estabelecido pelas metacomponentes); e as componentes de *aquisição de conhecimento* (processos envolvidos na aprendizagem dos modos de funcionamento das metacomponentes e das componentes de execução) (Sternberg, 1985, 1997b, 1999, 2003a).

Estas componentes de processamento de informação, entendidas como mais universais do que as aptidões na acepção tradicional, já que remetem para processos subjacentes à inteligência e não para conteúdos sobre os quais ela opera (Sternberg, 1999, 2003a), aplicam-se a tarefas e situações com as quais o indivíduo tem um grau variável de familiaridade, em função da experiência passada. Define-se assim um contínuo entre dois extremos, o da total novidade e o da automatização, ao longo do qual, de acordo com a *Subteoria Experiencial*, as componentes não são avaliáveis com igual eficácia, pelo que para medir a inteligência será necessário contemplar simultaneamente o funcionamento das componentes e o nível de experiência na tarefa a que se aplicam. Dois níveis de experiência são identificados como críticos na avaliação da inteligência: o da *novidade moderada* (não absoluta novidade, que exigiria do indivíduo recursos não disponíveis) e o da *automatização iminente* (não automatização adquirida, que representaria uma economia no funcionamento das componentes de processamento mental, não favorecendo a sua avaliação) (Sternberg, 1985, 1997b).

A *Subteoria Contextual* especifica as funções da inteligência: quando as componentes são aplicadas a tarefas e situações, mais familiares ou menos familiares, têm em vista o ajustamento do indivíduo ao meio. Este ajustamento pode ser alcançado por *adaptação* (o indivíduo modifica-se para se ajustar às exigências do meio), por *modelação* (o indivíduo modifica o meio para o ajustar às suas características) e/ou por *selecção* (o indivíduo selecciona um novo meio onde alcança maior congruência entre exigências e características pessoais). De acordo com esta subteoria, a inteligência define-se num contex-

to, sendo inadequada a noção de contínuo abstracto (factor *g*) que existe num vácuo. A diversidade das concepções e definições de inteligência, das mais implícitas (Sternberg, 1985b) às mais científicas (Thorndike et al., 1921.; Sternberg e Detterman, 1986), e ao longo da história da humanidade (Richardson, 1991), atesta a relevância do contexto social e cultural na delimitação e na definição do construto de inteligência.

A Teoria Triárquica assume, assim, que a inteligência se define pelo sucesso do ajustamento (*match*) do indivíduo ao contexto sócio-cultural, não pelo seu nível de “aptidão geral” (Sternberg, 2000). Este sucesso é em parte objectivamente avaliável, no quadro dos valores sociais e culturais dominantes num determinado lugar e momento histórico; mas é também, e sobretudo, subjectivamente vivenciado, em função das necessidades, dos valores e das metas pessoais do indivíduo. Emerge assim um novo conceito de inteligência que se define pela relação funcional do indivíduo em contexto – a Inteligência Funcional (*Successful Intelligence*): capacidade para alcançar sucesso na vida, de acordo com os critérios do indivíduo inserido num determinado contexto sócio-cultural (Sternberg, 1996, 1999, 2003a). Esta capacidade para atingir sucesso implica que o indivíduo capitalize as suas potencialidades e corrija ou compense os seus défices, através do equilíbrio entre três formas de inteligência: a *inteligência analítica* (que envolve as componentes de processamento de informação aplicadas a tarefas e situações relativamente familiares e de conteúdo abstracto, como a maioria das tarefas escolares), a *inteligência criativa* (que envolve as componentes de processamento de informação aplicadas a tarefas e situações relativamente novas) e a *inteligência prática* (que envolve as componentes de processamento de informação aplicadas a situações concretas do quotidiano, tendo em vista a adaptação, a modelação e/ou a selecção em relação ao meio) (Sternberg, 1999, 2003a).

Esta concepção é claramente contrastante com a perspectiva diferencial clássica da inteligência: primeiro, porque se baseia numa definição idiográfica (não nomotética) do sucesso (sucesso subjectivamente julgado, não apenas atribuído em função de critério externo); segundo, porque reconhece que a inteligência opera sempre num determinado contexto que lhe dá sentido funcional, sendo um conceito estéril se descontextualizado; terceiro, porque sublinha que a inteligência não só promove a adaptação ao meio, tomando-o como realidade que se impõe, mas pode implicar a modelação do próprio meio e/ou a selecção de um novo meio (adaptação como processo, activo e interactivo, não como produto da inteligência); quarto, porque é entendida no quadro de um sistema complexo que envolve diferentes elementos interrelacio-

nados (internos, externos, experienciais e relacionais), não como um conceito unitário que represente uma dimensão psicológica; quinto, porque a sua avaliação incide no equilíbrio entre diversas formas de inteligência, não numa quantificação expressa em termos de nível ou grau (Sternberg, 1997b, 1999, 2003a; Afonso, 1997, 2002, no prelo). Em suma, esta teoria apresenta-se congruente com uma psicologia positiva, ao entender a inteligência também no quadro da promoção do sucesso subjectivamente vivenciado ou bem-estar (Sternberg, 2003b) e contrasta por isso com a noção tradicional de inteligência que se desenvolveu a partir da abordagem “remediativa de Binet” – por alguns designada *mental orthopedics* (Sternberg, Torff e Grigorenko, 1998, p.374) – dirigida à correcção de défices identificados em dimensões cognitivas.

Implicações da Teoria Triárquica da Inteligência para a Educação e para a Avaliação da Inteligência

Dois tipos de estudos são descritos na literatura tendo em vista a validação da Teoria Triárquica da Inteligência: 1) os estudos de validação externa, geralmente envolvendo a aplicação da teoria ao planeamento e à avaliação de intervenções em contextos educativos; e 2) os estudos de validação interna, que pressupõem a medição da Inteligência Funcional e a análise da compatibilidade entre a estrutura das medidas e o modelo teórico subjacente.

1) Um dos propósitos do desenvolvimento da Teoria Triárquica da Inteligência foi, desde as suas primeiras formulações e de maneira manifesta, o de criar um quadro de referência conceptual que permitisse a articulação entre a avaliação das aptidões, o ensino e a avaliação dos conhecimentos. Tradicionalmente, esta ligação aptidões-instrução-avaliação não é nem completamente explícita nem solidamente fundamentada na teoria psicológica, algo que contribui para perpetuar um ciclo fechado (Sternberg, 1996, 1997a, 2000) em que as competências avaliadas pelos testes de inteligência geral, promovidas pela escola e valorizadas na avaliação dos alunos se validam mutuamente de um ponto de vista empírico, ainda que cubram apenas um espectro limitado do funcionamento cognitivo: o pensamento analítico e a memória. Por exemplo, as correlações entre os resultados dos testes de inteligência e as classificações escolares, que foram em geral interpretadas como evidência de validade empírica dos testes (Neisser, 1996), constituíram ao mesmo tempo evidência de validade de construto das avaliações

escolares, no quadro da docimologia (Miranda, 1981); mas, na ausência de uma conceptualização teórica explícita que reúna ambas as formas de avaliação reportando-as ao funcionamento da cognição humana, tais evidências empíricas resultam conceptualmente vazias e, por consequência, de duvidoso valor pragmático.

A Teoria Triárquica propõe-se exactamente constituir um tal quadro de referência conceptual, entre outros que o autor reconhece poderem servir o mesmo propósito (Sternberg, 1997a): por um lado, ao remeter para uma teoria psicológica, qualquer modelo psico-pedagógico que ligue aptidões-instrução-avaliação encontra fundamento, justificação e possibilidades de validação; por outro lado, ao ampliar a gama de funcionamento cognitivo abrangida pela avaliação da inteligência, pelo ensino e pela avaliação de conhecimentos, outras competências – para além da memória e do pensamento analítico, também a inteligência criativa e a inteligência prática – poderão vir a ser valorizadas pela escola e pela sociedade, aumentando o número dos que alcançam um grau satisfatório de adaptação ou sucesso, na escola e na vida quotidiana (Sternberg, 2002b; Sternberg e Clinkenbeard, 1995; Sternberg, Torff e Grigorenko, 1998).

Os estudos de validação externa da Teoria Triárquica da Inteligência no domínio da educação (ver Sternberg, 1999, 2002b, 2003; Sternberg e Grigorenko, 1997; Sternberg, Torff e Grigorenko, 1998) apoiam-se nesta perspectiva: a avaliação das aptidões, a instrução e a avaliação dos conhecimentos contemplam, cada uma, as três áreas de processamento mental identificadas pela teoria – as inteligências analítica, prática e criativa; os estudos envolvem uma de duas estratégias de inserção curricular, a de adição e a de infusão; e os programas de instrução, já implementados e em curso, abrangem uma gama variada de destinatários, desde o ensino básico ao universitário, e de conteúdos, desde a aprendizagem da leitura ao ensino da psicologia. Os resultados desta linha de investigação têm sido francamente favoráveis do ponto de vista da validação da teoria e promissores do ponto de vista da inovação psico-pedagógica: quando a instrução contempla áreas de processamento mental diferenciadas – a memória, a inteligência analítica, a inteligência prática e a inteligência criativa – mais estudantes encontram possibilidade de capitalizar as suas potencialidades e de corrigir e compensar os seus défices cognitivos, de onde decorrem níveis superiores de aproveitamento escolar e de satisfação pessoal (numa palavra, níveis superiores de “sucesso”). A ideia nuclear não é a de adequar os métodos e conteúdos do ensino às aptidões de cada estudante, facilitando-lhe a aprendizagem de modo artificial porque não representativo das situações

concretas do quotidiano, que geralmente exigem uma combinação das diversas áreas de processamento; a ideia é antes a de proporcionar a todos os estudantes oportunidades de desenvolvimento mais diversificadas, porque mais amplas no espectro do funcionamento cognitivo que abrangem, abrindo a mais estudantes a possibilidade de “sucesso” através de um desenvolvimento equilibrado, apoiado nas áreas de potencialidade e tendo em vista a correcção ou compensação das áreas de défice. Neste sentido, têm vindo a ser desenvolvidos princípios gerais e sugestões metodológicas para a implementação de práticas educativas formalmente fundamentadas na Teoria Triárquica da Inteligência (ver Sternberg, 1998a,b; Sternberg e Grigorenko, 1997).

2) A validação interna da teoria, em grande parte baseada em análises factoriais exploratórias e confirmatórias, pressupõe a construção de métodos de medida da Inteligência Funcional. O *Sternberg Triarchic Abilities Test (STAT)* (Sternberg, 1992, 2002c), foi o primeiro instrumento construído para a medição da inteligência funcional: assume o formato de um teste de papel-e-lápis, composto por nove partes de resposta por escolha múltipla – resultantes da aplicação das inteligências analítica, prática e criativa a três Áreas de Conteúdo, verbal, quantitativa e figurativa – e uma 10ª parte com três itens de resposta aberta – itens ensaio – um para cada aspecto da inteligência.

Os estudos de análise factorial confirmatória e exploratória relatados na literatura (Sternberg, 1999, 2003a, 2003c; Sternberg et al., 2000, 2001) têm fornecido resultados não completamente consistentes, quer ao longo do tempo, quer internacionalmente: ainda que em geral estes resultados, de acordo com o autor (Sternberg, 2003c), pareçam suportar em larga medida a teoria e a sua operacionalização – o *STAT*, actualmente na sua versão revista (2002c) – é de assinalar também que têm suscitado, juntamente com outras vertentes da teoria e da investigação empírica, acesa contestação por parte de alguns especialistas [ver, por exemplo, número especial da revista *Intelligence* (2003), 31(4)]. As dificuldades subjacentes à medição da inteligência criativa e da inteligência prática, por contraste com a inteligência analítica medida desde há muito pelos testes tradicionais de inteligência, decorrem em parte de se adequarem mal ao formato de resposta por escolha múltipla (Sternberg, 2003c); um formato de resposta aberta e de classificação dos produtos por juízes, de acordo com critérios pré-estabelecidos com base na teoria, tem vindo a ser por isso ensaiado nos estudos mais recentes de validação da Teoria Triárquica da Inteligência (ver, por exemplo, Grigorenko e Sternberg, 2001).

Estudos Portugueses com o *STAT-R* (H)

A versão revista do *STAT*, nível H (destinado a estudantes do ensino superior e a adultos) – *STAT-R* (H) (2002) – foi traduzida e adaptada para Portugal tendo em vista a utilização em investigação. A última versão portuguesa do teste (Afonso, 2004) retém apenas as nove partes constituídas por itens de resposta por escolha múltipla e contém um total de 45 itens, 5 por cada parte. A adaptação e revisão sucessiva, iniciada em 2002, comportou um estudo piloto com universitários, um primeiro ensaio experimental com uma amostra alargada de estudantes e adultos portugueses (Afonso, 2003) e um segundo ensaio experimental, do qual se apresentam de seguida alguns resultados.

A presente amostra é constituída por 370 participantes (199 F / 171 M), com idades compreendidas entre os 16 e os 70 anos (Mediana: 23 anos; 70% entre os 20 e os 29 anos) e escolarização entre os 6 e os 18 anos (Mediana: 12 anos de escolarização; 74% com 9 a 14 anos). A percentagem de estudantes na amostra é de 45% (n=165) e os restantes participantes concentram-se nas profissões intelectuais e científicas, técnicas e administrativas (36%). A amostra foi recolhida na sua maioria na região de Lisboa e Vale do Tejo (78%) embora apenas 9% dos participantes sejam residentes na cidade de Lisboa.

Neste trabalho optou-se por apresentar resultados significativos do ponto de vista conceptual, obtidos na análise factorial confirmatória: estes resultados permitem estabelecer relação entre as medidas obtidas no *STAT-R* e a Teoria Triárquica da Inteligência subjacente à sua construção. Ainda que tais resultados possam ser interpretados como evidências com implicações para a validação da teoria, e é esse o sentido que o próprio autor lhes tem atribuído em estudos semelhantes (Sternberg et al., 2000, 2001), está também em causa a validação do próprio método de medida enquanto operacionalização da teoria. Mais ainda, este tipo de estudo é essencial na adaptação de instrumentos de avaliação psicológica, como meio de averiguação do paralelismo entre as estruturas de resultados dos testes original e adaptado.

Quadro 1 - Estatísticas descritivas (média e desvio-padrão) e índices de consistência interna (alfa de Cronbach) para os Subtestes, as Áreas de Processamento Mental e as Áreas de Conteúdo do STAT-R (H) (N=370)

Estatísticas Descritivas									
Subtestes*	1 AV	2 AQ	3 AF	4 PV	5 PQ	6 PF	7 CV	8 CQ	9 CF
M	3.84	1.85	2.87	3.19	1.15	1.47	2.91	3.50	2.43
dp	1.01	1.26	1.36	1.02	1.13	1.16	1.08	1.46	1.34
Áreas de Processamento Mental	ANALÍTICA			PRÁTICA			CRIATIVA		
M	8.55			5.81			8.85		
dp	2.45			2.17			2.91		
Áreas de Conteúdo	VERBAL			QUANTITATIVA			FIGURATIVA		
M	9.94			6.49			6.77		
dp	2.14			2.83			2.83		
Consistência Interna (alfa de Cronbach)									
Subtestes*	1 AV	2 AQ	3 AF	4 PV	5 PQ	6 PF	7 CV	8 CQ	9 CF
	.08	.47	.46	.07	.44	.36	.21	.61	.43
Áreas de Processamento Mental	ANALÍTICA			PRÁTICA			CRIATIVA		
	.54			.43			.67		
Áreas de Conteúdo	VERBAL			QUANTITATIVA			FIGURATIVA		
	.42			.69			.63		
*Subtestes:									
AV: Analítica/Verbal			PV: Prática/Verbal			CV: Criativa/Verbal			
AQ: Analítica/Quantitativa			PQ: Prática/Quantitativa			CQ: Criativa/Quantitativa			
AF: Analítica/Figurativa			PF: Prática/Figurativa			CF: Criativa/Figurativa			

O QUADRO 1 apresenta algumas estatísticas descritivas dos resultados dos nove subtestes, das três Áreas de Processamento Mental de informação (partes ANALÍTICA, PRÁTICA e CRIATIVA), e das três Áreas de Conteúdo (partes VERBAL, QUANTITATIVA e FIGURATIVA). No mesmo quadro encontram-se os índices de consistência interna (alfa de Cronbach) que se situam entre .43 e .67 para as Áreas de Processamento Mental, entre .42 e .69 para as Áreas de Conteúdo e atingem valores entre .07 e .61 para os subtestes. Estes resultados, que encontram paralelo em estudos internacionais com o STAT original (1992) nos EUA, e com adaptações dessa versão do teste para outros países (Espanha e Finlândia) (Sternberg et al., 2000, 2001), justificam-se em parte pelo reduzido número de itens de cada subteste, apenas cinco, e também por alguma heterogeneidade do conteúdo das partes compostas que, ou contem-

plam diversidade de conteúdos dentro de uma Área de Processamento Mental, ou contemplam diversidade quanto às Áreas de Processamento Mental envolvidas em cada Área de Conteúdo.

Quadro 2 - Matriz de Intercorrelações dos Subtestes – STAT-R (H) (N=370)

Subtestes*	1 AV	2 AQ	3 AF	4 PV	5 PQ	6 PF	7 CV	8 CQ
2 AQ	.13							
3 AF	.14	.24						
4 PV	.18	.14	.23					
5 PQ	.05	.42	.29	.17				
6 PF	.06	.20	.19	.04	.21			
7 CV	.17	.27	.34	.27	.22	.18		
8 CQ	.22	.25	.40	.31	.26	.27	.32	
9 CF	.17	.18	.41	.25	.30	.29	.35	.34

*Subtestes:

AV: Analítica/Verbal

AQ: Analítica/Quantitativa

AF: Analítica/Figurativa

PV: Prática/Verbal

PQ: Prática/Quantitativa

PF: Prática/Figurativa

CV: Criativa/Verbal

CQ: Criativa/Quantitativa

CF: Criativa/Figurativa

Os coeficientes $\geq .10$ são significativos ao nível de significância de .05.

Os coeficientes $\geq .13$ são significativos ao nível de significância de .01.

A matriz de correlações entre os nove subtestes (QUADRO 2) apresenta coeficientes que se situam entre .04 e .42, sendo na maioria significativas ao nível de significância de .01. As correlações mais elevadas registam-se entre subtestes que envolvem um mesmo tipo de conteúdo – subtestes QUANTITATIVOS (ANALÍTICO e PRÁTICO) (.42) e subtestes FIGURATIVOS (ANALÍTICO e CRIATIVO) (.41) – mas também entre partes diferentes – entre os subtestes ANALÍTICO/FIGURATIVO e CRIATIVO/QUANTITATIVO (.40). As correlações dentro da parte CRIATIVA são mais homogêneas e elevadas (superiores a .30) do que as correlações dentro de cada uma das outras Áreas de Processamento.

Na análise factorial confirmatória, seguindo uma metodologia semelhante à de outros estudos internacionais (Sternberg et al., 2000, 2001), partiu-se da matriz de intercorrelações dos itens que se verificou previamente ser significativamente diferente da matriz identidade (teste de esfericidade de Bartlett: $\chi^2=2270.362$, $p<.000$, $gl=990$), podendo também a amostragem ser considerada adequada à aplicação de metodologia factorial (medida de Kaiser-Meyer-Olkin =.75). Através da aplicação do programa AMOS 5 (método da máxima verosimilhança), e assumindo a distribuição normal multivariada dos itens (médias dos índices de simetria e de curtose dos itens de -.12 e -.14, respectivamente), foi testado o ajustamento da estrutura das variáveis (itens) a sete modelos conceptuais:

- Modelo 1: Um factor geral (*g*) em que saturam todos os itens;
- Modelo 2: Três factores oblíquos correspondentes às Áreas de Processamento Mental – ANALÍTICA (itens 1 a 15, subtestes 1, 2 e 3), PRÁTICA (itens 16 a 30, subtestes 4, 5 e 6) e CRIATIVA (itens 31 a 45, subtestes 7, 8 e 9);
- Modelo 3: Três factores oblíquos correspondentes às Áreas de Conteúdo – VERBAL (itens dos subtestes 1, 4 e 7), QUANTITATIVA (itens dos subtestes 2, 5 e 8) e FIGURATIVA (itens dos subtestes 3, 6 e 9) – modelo multifactorial tradicional;
- Modelo 4: Nove factores ortogonais correspondentes aos nove subtestes;
- Modelo 5: Modelo hierárquico com nove factores de primeira ordem e um factor geral (*g*) de segunda ordem;
- Modelo 6: Modelo hierárquico com nove factores de primeira ordem e três factores oblíquos de segunda ordem, correspondentes às Áreas de Processamento Mental (ANALÍTICA, PRÁTICA e CRIATIVA) – modelo que representa a Teoria Triárquica;
- Modelo 7: Modelo hierárquico com nove factores de primeira ordem e três factores oblíquos de segunda ordem correspondentes às Áreas de Conteúdo (VERBAL, QUANTITATIVA e FIGURATIVA).

Quadro 3 - *Análise Factorial Confirmatória: Índices de ajustamento para os modelos testados (N=370)*

Índices										
Modelos	χ^2	gl	<i>p</i>	χ^2 /gl	GFI	PGFI	RMSEA	CFI	RFI	TLI*
Modelo 1 (<i>g</i>)	1301.38	945	.00	1.38	.86	.79	.032	.74	.43	--
Modelo 2 (APC)	1292.72	942	.00	1.37	.86	.79	.032	.75	.43	.03
Modelo 3 (VQF)	1265.81	942	.00	1.34	.87	.79	.031	.77	.44	.11
Modelo 4 (9 factores ortogonais)	1702.25	945	.00	1.80	.80	.73	.047	.45	.25	--
Modelo 5 (9 factores / <i>g</i>)	1162.19	936	.00	1.24	.88	.80	.026	.84	.48	.37
Modelo 6 (9 factores/ APC)	Não convergiu									
Modelo 7 (9 factores/ VQF)	1158.78	933	.00	1.24	.88	.79	.026	.84	.48	.37

Estatísticas de Ajustamento: χ^2 /gl (χ^2 / graus de liberdade); GFI (Goodness of Fit Index); PGFI (Parsimony Goodness of Fit Index); RMSEA (Root Mean Square Error of Approximation); CFI (Comparative Fit Index); RFI (Relative Fit Index); TLI (Tucker-Lewis Index)

* Comparação com o Modelo 1

No QUADRO 3 encontra-se um resumo das estatísticas de ajustamento relativas a cada modelo testado. O modelo que se mostra menos ajustado aos dados é o Modelo 4, que postula 9 factores independentes, com um índice CFI muito baixo (.45) e um valor χ^2/df relativamente elevado (1.80). O Modelo 1, que admite um único factor geral, para os mesmos graus de liberdade (945) apresenta índices de ajustamento mais favoráveis (CFI=.74 e $\chi^2/df=1.38$). Ainda assim, quer o Modelo 2 quer o Modelo 3, quando comparados com o Modelo 1 através do índice TLI, algo possível por serem nele embutidos, mostram incremento, ainda que muito ligeiro, respectivamente, .03 e .11. No Modelo 2 as correlações entre os factores, deixadas como parâmetros livres, são de .89 (Analítico-Criativo), .90 (Prático-Criativo) e .93 (Analítico-Prático) e no Modelo 3 são de .78 (Verbal-Quantitativo), .82 (Quantitativo-Figurativo) e .87 (Verbal-Figurativo). Os modelos 5 e 7 apresentam índices de ajustamento, e índices de incremento relativamente ao Modelo 1 (TLI=.37), notavelmente coincidentes, o que os torna equivalentes em termos de ajustamento à estrutura empírica dos dados e o Modelo 6 não convergiu, apesar do elevado número de iterações ensaiado (500). Os índices RMSEA mais próximos de 0 e os índices GFI mais próximos de 1 dos modelos 5 e 7 permitem considerá-los mais ajustados aos dados do que qualquer dos modelos de primeira ordem ou multifactoriais.

Embora no conjunto os resultados pareçam de certo modo favoráveis à consideração de uma estrutura hierárquica para os itens do *STAT-R*, a estrutura parece não assumir nesta amostra uma configuração concordante com a previsível à luz da Teoria Triárquica da Inteligência. Há aliás a assinalar que o modelo que assume os três factores postulados pela Teoria Triárquica (Modelo 6) não convergiu nesta amostra, à semelhança do que aconteceu no primeiro ensaio experimental (Afonso, 2005b). Também os estudos internacionais com diversas versões do *STAT* têm apresentado resultados pouco consistentes que, embora apontem por vezes no sentido de uma ligeira vantagem do Modelo Triárquico na descrição da estrutura dos dados, demonstram ainda assim a existência de modelos alternativos com níveis de ajustamento aos dados muito aproximados (Sternberg et al., 2000,2001).

Conclusão

A Teoria Triárquica da Inteligência propõe um alargamento da perspectiva diferencial de investigação e de avaliação da inteligência com consequências significativas no plano das aplicações, designadamente no domínio educativo.

Tratando-se de uma teoria sistêmica, que define a inteligência contextualmente, não só amplia a gama de aptidões tomadas como pertinentes para o ajustamento ao meio, como articula níveis diferenciados de observação e análise da cognição humana (na acepção de Reuchlin, 1999/2002), desde o mais molecular (componentes de processamento da informação) ao mais molar (o sucesso adaptativo do comportamento em contexto). Deste modo, mais do que contestar a perspectiva tradicional, esta teoria parece contextualizá-la, oferecendo um quadro de referência mais integrado e compreensivo, donde mais próximo da realidade complexa do funcionamento cognitivo; parafraseando um título de Sternberg (2002a), *intelligence is not just inside the head* [a inteligência não está simplesmente na cabeça].

Apesar do enorme valor heurístico e hermenêutico da teoria, a sua operacionalização em métodos de medida parece ainda francamente limitada nos planos metrológico e conceptual. Uma diversidade de estudos de validação interna da teoria tem vindo a demonstrar a dificuldade de construção de medidas da inteligência prática e da inteligência criativa, sobretudo enquanto se persiste em adoptar um formato semelhante ao dos testes tradicionais, com itens de resposta por escolha múltipla e de cotação dicotómica. A investigação ganhará em progredir no sentido de ampliar mais do que apenas a teoria: também a concepção subjacente à medida cognitiva, o formato dos instrumentos e as técnicas de construção e investigação metrológica terão de evoluir, tomando por referência a predição do sucesso subjectivo ou bem-estar, para lá do sucesso objectivo ou rendimento, e aproximando-se assim, nalguns aspectos, da avaliação das condutas conativas.

Retomemos por fim a pergunta inicial: Charles Darwin e Albert Einstein eram homens inteligentes? Se admitirmos como parte da definição de inteligência o sucesso subjectivo, a vivência íntima do ajustamento e do bem-estar, a resposta deveria ser em parte deles próprios. Podemos acreditar que, como qualquer ser humano, ambos tenham enfrentado momentos de desalento e desadaptação; mas o que fez o seu sucesso foi a forma notável como ambos os ultrapassaram e evoluíram, através de uma gestão eficaz dos seus recursos e oportunidades e baseando-se no reconhecimento das suas potencialidades e défices no plano cognitivo – nas palavras (de curiosa antecipação conceptual) de Einstein, nos planos do “pensamento abstracto”, da “imaginação” e da “capacidade prática” (cf. Pág. 339). Mais do que as obras que nos deixaram, testemunhos inquestionáveis do seu sucesso objectivo, foi a maneira como as construíram que terá feito deles homens inteligentes.

Referências

- Afonso, M.J. (1997, Dezembro). “*Successful Intelligence*”: um novo conceito, uma perspectiva abrangente de medida da inteligência. Comunicação oral apresentada na V Conferência Internacional de Avaliação Psicológica. Braga, Portugal.
- Afonso, M.J. (2002). Inteligência Funcional: aspectos heurísticos e hermenêuticos do construto. *Revista Portuguesa de Psicologia*, 36, pp. 9-24.
- Afonso, M.J. (2003, Outubro). *STAT-R: Teste Triárquico de Aptidões de Sternberg – Revisto (nível H). Estudo Piloto e Primeiro Ensaio Experimental na População Portuguesa*. Comunicação Poster ao V Simpósio Nacional de Investigação em Psicologia, Lisboa, Associação Portuguesa de Psicologia.
- Afonso, M.J. (2004). *Teste Triárquico de Aptidões de Sternberg – Forma Portuguesa, Revisita. (STAT-R, H, 2002)*. Lisboa: Centro de Psicometria e Psicologia da Educação.
- Afonso, M.J. (2005a). Challenging Practices: Implications of “Successful Intelligence” for Career Counseling and Guidance. In *Actas da Conferência Internacional da AIOSP/IAEVG 2005*. Portugal: IOP/Universidade de Lisboa.
- Afonso, M.J. (2005b, Outubro). A “*Inteligência Funcional*” e a Amostragem de Funções: a versão portuguesa do *STAT-R (H) (2004) no Panorama Internacional*. Comunicação oral apresentada no I Simpósio Internacional de Inteligência Humana – Investigação e Aplicações, Universidade de Évora.
- Carroll, J. (1982). The measurement of intelligence. In R.J. Sternberg (Ed.), *Handbook of Human Intelligence* (pp. 29-120). Cambridge: Cambridge University Press.
- Darwin, F. (Ed) (1905). *The Life and Letters of Charles Darwin*. New York: Appleton.
- Detterman, D.K. (Ed.). (2003). (Sem título). *Intelligence*, 31 (4) (Número Especial).
- Gardner, H. (1983). *Frames of Mind. The Theory of Multiple Intelligences*. New York, NY: Basic Books.
- Gardner, H. (1999). *Intelligence Reframed: Multiple Intelligences for the 21st Century*. New York, NY: Basic Books.
- Grigorenko, E.L. e Sternberg, R.J. (2001). Analytical, creative, and practical intelligence as predictors of self-reported adaptive functioning: a case study in Russia. *Intelligence*, 29, pp. 57-73.
- Instituto António Houaiss (2004). *Dicionário Eletrônico Houaiss da Língua Portuguesa* (versão 1.05). Brasil: Editora Objectiva Lda.
- Miranda, M.J. (1981). A docimologia em Perspectiva. *Revista Portuguesa de Pedagogia*, XV, pp. 75-114.
- Miranda, M.J. (1986). Perspectivas da Investigação e Avaliação da Inteligência. *Revista Portuguesa de Psicologia*, 23, pp. 27-54.
- Miranda, M.J. (2000). La investigación psicológica de la inteligencia humana: balance del siglo en la transición del milenio. *Persona*, 3, pp. 27-45.

- Miranda, M.J. (2002). A Inteligência Humana: Contornos da Pesquisa. *Paidéia*, 12 (23), pp. 19-29.
- Miranda, M.J. (2004, Agosto). *The Psychology of Human Intelligence: the State of the Art*. Comunicação poster apresentada no 28º Congresso Internacional de Psicologia, Beijing, China. (Resumo: *International Journal of Psychology*, 2004)
- Neisser, U.; Boodoo, G.; Bouchard, T.; Boykin, W.; Brody, N.; Ceci, S. et al. (1996). Intelligence: Knowns and Unknowns. *American Psychologist*, 51 (2), pp.77-101.
- Nyborg, H. (Ed.) (2003). *The Scientific Study of General Intelligence. Tribute to Arthur Jensen*. Oxford: Pergamon.
- Pais, A. (2004). “*Subtil é o Senhor*” – *Vida e Pensamento de Albert Einstein*. (F. Parente e V. Esteves trad.). Lisboa: Gradiva. (Obra original publicada em 1982).
- Resnick, L. e Resnick, D.P. (1982). Testing in America: The current challenge. *International Review of Applied Psychology*, 31 (Sp.Issue), pp. 7-34.
- Reuchlin, M. (2002). *Evolução da Psicologia Diferencial*. (M.J. Miranda e M.J. Afonso trad.). Lisboa: Fundação Calouste Gulbenkian (Obra original publicada em 1999).
- Richardson, K. (1991). *Understanding Intelligence*. Philadelphia: Open University Press.
- Sternberg, R.J. (1982). Conceptions of Intelligence. In R.J. Sternberg (Ed.), *Handbook of Human Intelligence* (pp.3-28). Cambridge: Cambridge University Press.
- Sternberg, R.J. (1985a). *Beyond IQ. A Triarchic Theory of Human Intelligence*. Cambridge: Cambridge University Press.
- Sternberg, R.J. (1985b). Implicit Theories of Intelligence, Creativity, and Wisdom. *Journal of Personality and Social Psychology*, 49, pp. 607-627.
- Sternberg, R.J. (1990). *Metaphors of Mind. Conceptions of the Nature of Intelligence*. Cambridge: Cambridge University Press.
- Sternberg, R.J. (1992). *Sternberg Triarchic Abilities Test (STAT)*. (Teste não publicado).
- Sternberg, R.J. (1996). *Successful Intelligence. How Practical and Creative Intelligence Determine Success in Life*. New York, NY: Simon & Schuster.
- Sternberg, R.J. (1997a). Educating Intelligence: Infusing the Triarchic Theory into School Instruction. In R.J. Sternberg e E.L. Grigorenko (Eds.), *Intelligence, Heredity and Environment* (pp. 343-362). New York, NY: Cambridge University Press.
- Sternberg, R.J. (1997b). The Triarchic Theory of Intelligence. In D.P. Flanagan e J.L. Gensholf (Eds.), *Contemporary Intellectual Assessment: Theories, Tests and Issues* (pp. 92-104). New York, NY: Guilford Press.
- Sternberg, R.J. (1998a). Applying the Triarchic Theory of Human Intelligence in the Classroom. In R.J. Sternberg, e W.M. Williams (Eds.), *Intelligence, Instruction and assessment: theory into practice* (pp. 1-15). Mahwah, NJ: Lawrence Erlbaum Associates.
- Sternberg, R.J. (1998b). Principles of Teaching for Successful Intelligence. *Educational Psychologist*, 33 (2/3), pp. 65-72.

- Sternberg, R.J. (1999). The Theory of Successful Intelligence. *Review of General Psychology*, 3 (4), pp. 292-316.
- Sternberg, R.J. (2000). Implicit Theories of Intelligence as Exemplar Stories of Success: Why Intelligence Test Validity is in the Eye of the Beholder. *Psychology, Public Policy, and Law*, 6 (1), pp.159-167.
- Sternberg, R.J. (2002a). Intelligence is not just Inside the Head: The Theory of Successful Intelligence. In J. Aronson (Ed.), *Improving Academic Achievement: Impact of Psychological Factors on Education* (pp. 227-244). San Diego, CA: Academic Press.
- Sternberg, R.J. (2002b). Raising the Achievement of All Students: Teaching for Successful Intelligence. *Educational Psychology Review*, 14 (4), pp. 383-393.
- Sternberg, R.J. (2002c). *Sternberg Triarchic Abilities Test - Revised (STAT-R)*. (Teste não publicado).
- Sternberg, R.J. (2003a). A Broad View of Intelligence: The Theory of Successful Intelligence. *Consulting Psychology Journal*, 55 (3), pp.139-154.
- Sternberg, R.J. (2003b). Driven to Despair: Why we need to redefine the concept and measurement of intelligence. In L.G. Aspinwall e U.M. Stander (Eds.), *A Psychology of Human Strengths: Fundamental questions and future directions for positive psychology*. Washington, DC: American Psychological Association.
- Sternberg, R.J. (2003c). Our Research Program Validating the Triarchic Theory of Successful Intelligence: Reply to Gottfredson. *Intelligence*, 31 (4), pp. 399-413.
- Sternberg, R.J., Castejón, J.L., Prieto, M.D., Hautamäki, J. e Grigorenko, E.L. (2001). Confirmatory Factor Analysis of the STAT in Three International Samples: An Empirical Test of the Triarchic Theory of Intelligence. *European Journal of Psychological Assessment*, 17 (1), pp. 1-16.
- Sternberg, R.J. e Clindenbeard, P.R. (1995). The Triarchic Model Applied to Identifying, Teaching, and Assessing Gifted Children. *Roeper Review*, 17 (4), pp. 255-260.
- Sternberg, R. J. e Detterman. D. K. (Eds.) (1986). *What is intelligence?* Norwood, NJ: Ablex.
- Sternberg, R.J., Prieto, M.D. e Castejón, J.L. (2000). Análisis Factorial Confirmatorio del Sternberg Triarchic Abilities Test (nivel-H) en una muestra española: resultados preliminares. *Psicothema*, 12 (4), pp. 642-647.
- Sternberg, R.J., Torff, B. e Grigorenko, E.L. (1998). Teaching Triarchically Improves School Achievement. *Journal of Educational Psychology*, 90 (3), pp. 374-384.
- Taylor, F.W. (1911). *The Principles of Scientific Management*. New York, NY: Harper & Row.
- Thorndike, E.L. et al. (1921). Intelligence and its measurement. *Journal of Educational Psychology*, 12.
- Wechsler, D. (1944). *The measurement of adult intelligence* (3rd ed.). Baltimore: Williams and Wilkins.

INTELLIGENCE, EDUCATION AND SUCCESS: THE DIFFERENTIAL AND THE SYSTEMIC APPROACHES TO HUMAN INTELLIGENCE AND THEIR IMPLICATIONS FOR EDUCATION

Maria João Afonso

Faculdade de Psicologia e de Ciências da Educação, Universidade de Lisboa, Portugal

Abstract: This paper deals with the notion of success and its relation to intelligence, within two paradigms of research and assessment: the differential paradigm, which is behind the traditional intelligence testing, and the systemic paradigm, here represented by one of the most influential current theories, the Sternberg's Triarchic Theory of Human Intelligence. Some implications of the theory and of the embedded concept of Successful Intelligence to education and intelligence assessment are analyzed. The Sternberg Triarchic Abilities Test – Revised (STAT-R) is presented as an operationalization of the theory, and some results obtained in Portugal, with the Portuguese experimental version (level H – college students and adults) (Afonso, 2004) are discussed: internal consistency and confirmatory factor analysis.

KEY-WORDS: *Intelligence, successful intelligence, STAT-R (H), confirmatory factor analysis.*

COMPETÊNCIA SOCIAL: A SUA AVALIAÇÃO EM CONTEXTOS DE DESENVOLVIMENTO E EDUCAÇÃO

Adelinda Araújo Candeias
(Universidade de Évora)

Leandro S. Almeida
(Universidade do Minho)

Resumo

A competência social, como um dos domínios da inteligência prática, assume nos nossos dias uma importância crescente no estudo das aptidões e do funcionamento cognitivo. A sua definição inclui aspectos cognitivos e contextuais, aspectos que devem ser devidamente ponderados nos testes psicológicos propostos para a sua avaliação. Este artigo descreve uma escala portuguesa para a avaliação da competência social assente nas percepções pessoais de competência e nas nomeações de colegas e professores para a resolução de problemas sociais específicos. Uma amostra de 326 alunos entre o 7º e o 11º anos de escolaridade foi usada para apreciar a consistência interna dos itens e a sua estrutura factorial. Os resultados apontam para um factor geral de competência social, acrescido de outros cinco factores reflectindo as especificidades das situações sociais representadas nos itens (intimidade, liderança, relacionamento com os pares, relacionamento com adultos, e relacionamentos com os pais). Os coeficientes de correlação entre a competência social e as classificações escolares dos alunos aproximaram-se de zero, sendo positivas e significativas as correlações com as nomeações pelos colegas e pelos professores.

PALAVRAS-CHAVE: *Competência social; inteligência; resolução de problemas sociais; avaliação cognitiva.*

Introdução

A compreensão da competência humana é um dos desafios mais presentes na história da Psicologia. As abordagens atomistas, dominantes durante grande parte do século XX, desempenharam um papel importante na identificação e caracterização dos diversos elementos inerentes a este conceito. Reportando-nos à competência social, tais estudos tiveram o mérito de progressivamente identificarem os diversos *skills* subjacentes (Spivak, Platt & Shure, 1976; Shure & Spivak, 1988; Oppenheimer, 1989; Stricker & Rock,

Morada (address): Adelinda Candeias, Departamento de Psicologia, Universidade de Évora, 7001 Évora – Portugal;
E-mail: aac@uevora.pt

1990; Pelechano, 1994), os seus diversos níveis de desenvolvimento estrutural (Selman, 1976, 1980; Turiel, 1980, 1983, 1994; Gibbs & Widaman, 1982; Yeates & Selman, 1989) ou, ainda, os elementos externos de reconhecimento da eficácia comportamental pelo próprio e pelos outros (Ford, 1982; Ford & Tisak, 1983). Como noutros domínios da Psicologia, temos assistido nas últimas duas décadas a uma progressiva integração desses elementos em modelos mais holísticos e sistémicos com o objectivo de representar a complexidade e o dinamismo deste conceito, numa abertura progressiva às variáveis contextuais no desenvolvimento e desempenho cognitivos (Gardner, 1983, 1999; Greenspan & Driscoll, 1997; Almeida & Roazzi, 1988; Mettrau & Almeida, 1996; Candeias & Almeida, 1999; Sternberg & Grigorenko, 2003).

Um dos modelos holísticos na área é representado por Greenspan e colaboradores. Greenspan e Driscoll (1997) propõem um modelo conceptual da competência humana partindo de trabalhos teóricos e empíricos de outros autores (e.g., Gardner, 1983, 1999; Guilford, 1959, 1985). O modelo estrutura-se em três níveis, de um mais global (competência pessoal) até um mais molecular ou elementar (inteligência social e competência social). Como referem os autores *"as in any classificatory scheme, groupings of elements are somewhat arbitrary and can be changed depending on whether one focuses on 'intensional' or 'extensional' (or in this case, intellectual or social) features"* (Greenspan & Driscoll, 1997, p.133).

Neste modelo, o 'espectro da competência humana' (*Ibidem*, p.132) é explicado a partir dos conceitos de *inteligência* e de *competência pessoal* que aparecem como estruturantes da capacidade do indivíduo resolver tarefas e desafios específicos. O conceito de competência pessoal tem um significado amplo, incorporando todos os *skills* que contribuem para realizar objectivos ou para resolver desafios, enquanto o termo inteligência tem um significado mais próximo do conjunto de *skills* envolvidos no processo de pensar e compreender (Greenspan & Driscoll, 1997). Assim, a inteligência constitui um pré-requisito para se ser bem sucedido em qualquer tipo de tarefa, apesar de coexistir com outros aspectos da competência que contribuem para a realização das tarefas com sucesso.

A competência pessoal subdivide-se em quatro 'entidades globais' ou domínios de competência: física, afectiva, quotidiana e académica. Estas, por sua vez, dividem-se em subcomponentes, sendo a *inteligência social* um dos subcomponentes da competência quotidiana, a par da *inteligência prática* (*Ibidem*, p. 133). Enquanto a primeira se refere à capacidade de pensar e compreender problemas interpessoais, a segunda refere-se a problemas

do dia-a-dia, de natureza técnica, mecânica e física. Neste contexto, a inteligência social constitui um subdomínio de natureza cognitiva que intersecta, simultaneamente, os domínios da competência social, intelectual e quotidiana.

Anteriormente, Greenspan (1979) tinha considerado a inteligência social como um dos domínios da inteligência adaptativa, a par da inteligência prática e da inteligência conceptual. Em 1981, numa remodelação do modelo, substitui inteligência adaptativa por competência adaptativa, mantendo os três domínios anteriores. A inteligência social aparece, então, com um dos subdomínios, em paralelo com o temperamento e o carácter, que integram a competência social e a competência intelectual. Nesta remodelação, Greenspan propõe que a inteligência social agrupa a *social sensitivity*, que se refere à capacidade para ler pistas sociais em situações interpessoais e que se subdivide em *inferência social* e *role-taking*, e a *social insight*, ou a capacidade para compreender os processos subjacentes à interacção social. Em síntese, com a versão do modelo de competência social apresentada em 1997, toma forma a hipótese, já antes avançada, de apontar a inteligência social como um subdomínio de natureza cognitiva, de intersecção entre *competência social*, *intelectual* e *quotidiana*, abrindo possibilidades de análise mais amplas da *competência pessoal* e da interacção entre inteligência e personalidade (Greenspan & Driscoll, 1997; Cantor & Kihlstrom, 1987, 1989; Kihlstrom & Cantor, 2000).

Contudo subsistem algumas dificuldades com o construto de inteligência social. Este facto pode dever-se à relativa ausência de itens de inteligência social nas escalas de avaliação utilizadas, verificando-se uma maior incidência dos itens nos componentes não intelectivos da competência social, com ênfase sobretudo nos componentes emocionais (Greenspan & Driscoll, 1997). Os cinco domínios que Gresham e Elliot (1990) apresentam para definir *skills* comportamentais sociais, por exemplo, estão todos relacionados com a competência afectiva: cooperação, assertividade, responsabilidade, empatia e autocontrolo. Os domínios de competência tomados não são suficientes para tirar conclusões acerca da validade do modelo a nível da inteligência social, pois são esquecidos os aspectos relacionados com os processos cognitivos. Alguns estudos em que se acrescentaram conteúdos suplementares de inteligência social às escalas de avaliação da competência (Mathias & Nettelbeck, 1992) vieram comprovar, através da análise factorial confirmatória, a validade e utilidade do construto de inteligência social.

Esta área emergente da inteligência traz alguma novidade à investigação, retomando a distinção entre inteligência quotidiana e académica, ou entre in-

teligência prática e social. A proposta de definição de inteligência social valoriza a distinção com aspectos mais afectivos e personalísticos (temperamento e carácter). Mesmo assim, todo este movimento de justaposição e de intersecção de conceitos gerais, como inteligência e personalidade, e de conceitos mais específicos como competência intelectual, competência social e competência quotidiana, apesar do seu irrefutável valor heurístico, coloca grandes desafios em termos da clarificação de cada conceito e da sua operacionalização. Muito concretamente, o conceito de inteligência social aparece aqui numa intersecção entre competência quotidiana, social e intelectual que importa clarificar no futuro, deixando em aberto a relação entre inteligência social e as competências que intersecta.

O conceito de competência qualifica os comportamentos em termos de relevância e eficácia comportamental. Um comportamento socialmente inteligente envolve processos psicológicos e contextuais, tais como: (i) *processos motivacionais*, incluem interesse e preocupação em realizar objectivos (Ford, 1986), as *crenças pessoais* relativas às próprias capacidades (Bandura, 1986), e a *expressividade emocional*, ou seja, as reacções aos sucessos e/ou fracassos sociais constituem um importante factor para desencadear ou não esforços para se ser socialmente inteligente (Ford, 1995); (ii) *processos relacionados com as capacidades (skill-related processes)*, incluem o conhecimento social sobre comportamentos, a capacidade de codificação de informação social, a capacidade de planeamento e de resolução de problemas sociais enquanto preditores do comportamento social eficaz (Ford, 1982; Spivak *et al.*, 1976); e (iii) *processos contextuais*, que integram a congruência entre o contexto social, os objectivos e as capacidades pessoais, incluindo ainda os recursos informativos e materiais e um clima emocional positivo (Ford, 1995). Na conjugação de todos estes processos, podemos assumir o modelo da “experiência em desenvolvimento” (Sternberg & Grigorenko, 2003) como uma importante ferramenta conceptual para compreendermos o conceito de competência humana. Os autores propõem o conceito de *experiência em desenvolvimento*, para se referirem ao *processo contínuo de desenvolvimento de experiência quando o indivíduo opera num determinado âmbito* (p.19). Assume-se que esse desenvolvimento potencial ocorre continuamente e está dependente da operação de um conjunto de elementos (Figura 1).

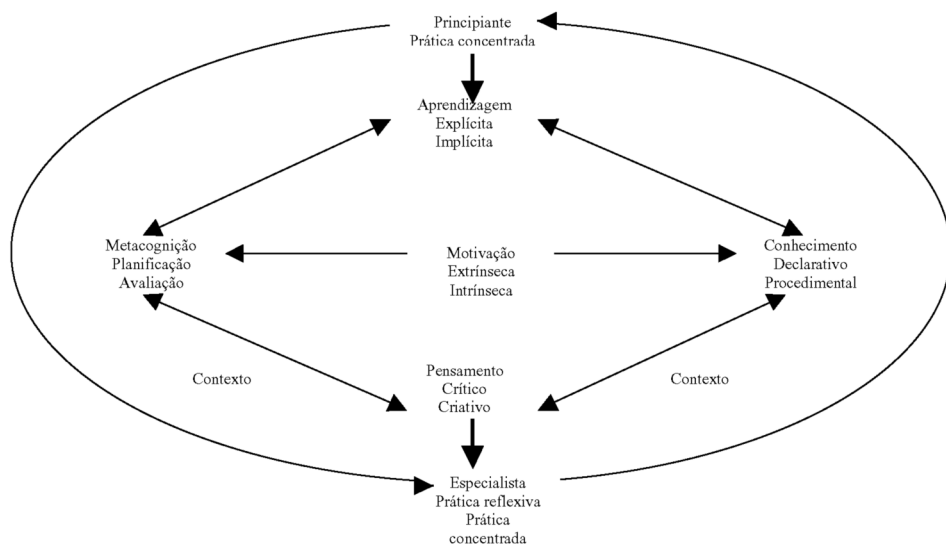


Figura 1 - Modelo de "experiência em desenvolvimento" (in Sternberg & Grigorenko, 2003, p.20)

Como podemos observar na figura 1, este modelo comporta seis elementos essenciais: as aptidões metacognitivas, as aptidões para a aprendizagem, o pensamento, o conhecimento, a motivação e a experiência. Estes elementos, embora distintos, interagem entre si. Os autores propõem as seguintes explicitação para cada um dos elementos do modelo (Sternberg & Grigorenko, 2003, pp. 21-3): *Aptidões metacognitivas* referem-se à compreensão e controlo que a pessoa tem sobre a sua própria cognição, por exemplo reconhecimento, definição e representação de problemas, formulação de estratégias, levantamento de recursos, supervisão e avaliação da resolução de problemas; *Aptidões para a aprendizagem* associadas à aquisição de informações e construção de conhecimentos; *Aptidões para o pensamento* ou conjuntos de aptidões que operam conjuntamente para a tradução do pensamento em ações no mundo real, por exemplo processos analíticos, criativos e práticos; *Conhecimento* que inclui o conhecimento declarativo referente a princípios, conceitos e leis, e o conhecimento procedimental relativo a procedimentos e estratégias; *Motivação* onde se pode incluir a motivação para o sucesso relativa à busca de desafios e riscos moderados e ao desenvolvimento de esforço para melhorar o desempenho e superar-se, e a motivação orientada para a competência ou para a eficácia pessoal, ou seja, crenças da pessoa sobre a sua própria capacidade para resolver e fazer face aos problemas (Bandura, 1986); *Contexto* tomando-se aqui factores contextuais que podem potenciar

ou dificultar a expressão e o desenvolvimento dos atributos do indivíduo, por exemplo a língua materna, a ênfase na rapidez, a importância que tem para o sujeito a tarefa que lhe é proposta ou a familiaridade com o material e o formato da mesma.

Colocando em funcionamento este modelo, dir-se-á que o aprendiz tenta adquirir competência mediante uma prática deliberada. Esta prática exige a interação entre os seis elementos do modelo. No centro e dirigindo esses elementos encontra-se a motivação, esta dirige as aptidões metacognitivas, que activam as aptidões para a aprendizagem e para o pensamento, permitindo que aumente o nível de experiência. Por sua vez, o conhecimento declarativo e procedimental, adquirido através da extensão das aptidões para pensar e aprender, faz com que essas aptidões se usem cada vez mais e com maior eficácia. Todos estes processos ocorrem num contexto e, simultaneamente, podem influenciá-lo.

Neste estudo, assumimos, então, que a competência humana é um conceito estruturante para a capacidade de realizar os objectivos valorizados pelo indivíduo através da resolução de tarefas ou desafios específicos. O conceito de competência pessoal tem um significado amplo, integrando processos inerentes ao indivíduo (metacognição, aprendizagem, motivação, pensamento e conhecimento) e processos inerentes ao contexto (adaptação e congruência entre os objectivos do sujeito, as suas capacidades e os desafios das situações, e os recursos informativos e materiais disponíveis no contexto). Cada um destes elementos é mobilizado em função dos níveis de *motivação* intrínseca e extrínseca do indivíduo para aquele tipo de situação que, por sua vez, activam os recursos metacognitivos. Os *recursos metacognitivos*, nomeadamente as habilidades para avaliar a situação e planear como actuar, regulam o desencadear de processos de aquisição de mais informação através da *aprendizagem* implícita ou explícita ou da mobilização de *conhecimentos* (declarativos e procedimentais) armazenados na memória. Os conteúdos adquiridos ou recuperados da memória serão analisados, seleccionados, combinados, re(criados) e aplicados, conforme as operações de *pensamento* mobilizadas: *crítico* (analítico) que inclui analisar, criticar, julgar, avaliar, comparar e contrastar, *criativo* que inclui criar, descobrir, inventar, imaginar e gerar hipóteses, e *prático* que inclui aplicar, usar e praticar. Por último, estes elementos serão mobilizados em função do tipo de informação das tarefas e que organizamos em competência pessoal, competência social e competência prática ou quotidiana (onde integramos competência académica e a competência profissional). A *competência pessoal* incidirá sobre conteúdos relativos ao próprio indivíduo (emoções e interesses, pensamentos e comportamentos); a *competência social* inci-

de sobre informações relativas aos outros indivíduos (suas emoções e interesses, seus pensamentos e comportamentos, as interações entre eles e com o próprio sujeito); e a *competência quotidiana ou prática* incidirá sobre conteúdos informativos formais e informais, ou organizados e ocasionais (conteúdos académicos e profissionais ou higiene e alimentação, por exemplo, respectivamente).

De acordo com este modelo, cada um dos elementos do processo mediacional é activado e, por sua vez, activa outro elemento formando uma complexa *rede operativa*. Esta proposta de conceptualização da competência humana requer uma operacionalização do conceito que tenha presente quer os elementos operativos quer os elementos informativos. Por outro lado, os processos mediacionais que caracterizam a interação entre indivíduo-tarefa-situação implicam uma avaliação destes sub-sistemas quer do ponto de vista da competência auto-percebida, da competência hetero-percebida e do desempenho (no sentido de *performance*) propriamente dito.

Face ao conceito de competência proposto é crucial procurar desenvolver provas que o operacionalizem, assim como testar a respectiva validade. Neste estudo, apresentamos uma prova de competência social de natureza situacional. Trata-se de uma prova cujos estudos preliminares de adaptação e estudo psicométrico decorreram em 2001 (Candeias, 2001), tomando as auto-avaliações e hetero-avaliações de competência comportamental em situações específicas do quotidiano social e interpessoal.

Método

Amostra

A amostra é constituída por 326 jovens do 7º (n=106), 9º (n=107) e 11º (n=113) anos de escolaridade das regiões norte (47,5%) e sul (52,5%) de Portugal. A maioria dos alunos pertencia a um estrato social médio (63,2%), havendo 23% reportados a um estrato social baixo e 13,8% ao estrato elevado. Questionados sobre o grau de sua participação social (actividades diversas), quase metade dos sujeitos menciona não terem qualquer participação (46%), havendo 42% que menciona participar e 12% que menciona participar muito. As idades oscilaram entre 11 e 20 anos (média de 14,7 e desvio-padrão de 1,84).

Instrumento

Podemos dizer que o desenvolvimento desta prova teve como referência o *Social Competence Nomination Form (SCNF – Ford, 1982)*, considerado um dos testes de avaliação mais fiáveis e válidos no âmbito das competências sociais. Na adaptação e desenvolvimento da prova para Portugal assumiu-se a designação “Teste de Competência Situacional” pois que a avaliação recorre a estímulos situacionais. Esta prova foi construída com o objectivo que cada situação proporcionasse referências comportamentais e contextuais específicas e detalhadas, podendo recorrer a percepções de pares, de professores e do próprio, a respeito da competência situacional do indivíduo. O teste é composto por seis situações sociais hipotéticas que requerem uma grande variedade de skills comportamentais e cognitivas (Quadro I).

Quadro 1 - Lista das situações-estímulo do SCNF

1. Situação de trabalho de casa

Na tua escola todos os alunos se queixam porque este ano os professores passaram trabalhos de casa para as férias do Natal. Os alunos da tua turma mais do que queixarem-se, juntaram-se e pediram aos professores para ouvirem o seu lado da história. Um grupo de professores concordou em conversar com um representante dos alunos sobre as suas razões, na próxima reunião de professores.

2. Situação de um encontro duplo

Estás bastante contente porque finalmente conseguiste um encontro com alguém de quem gostas à muito tempo. Contudo, foi-te pedido que fizesses dessa situação um encontro duplo, pois a pessoa com quem vais sair tem um(a) primo(a) de visita, em sua casa, durante esta semana, e não o pode deixar só. De facto, tens de encontrar alguém que acompanhe o(a) primo(a) da pessoa com quem vais sair. Assim, tens de pensar em alguém que seja bom conversador e que seja suficientemente perspicaz para saber quando se deverá ir embora e convidar o(a) primo(a) a ir consigo.

3. Situação de um estudante-visitante

Um dos teus melhores professores faleceu tragicamente num acidente. Os estudantes da tua classe juntaram-se e decidiram fazer algo pela família do professor. A turma decidiu que alguém deveria fazer uma visita pessoal à família do professor. Essa pessoa deverá levar flores e tentar expressar à família o quanto os estudantes lamentam a perda de um professor tão bom e tão amigo.

4. Situação de visita dos pais à escola

Um dos teus pais ou encarregado de educação (o pai ou a mãe) virá visitar a escola durante um dia no âmbito de um novo programa. É suposto que este programa deverá

permitir aos pais um melhor conhecimento do tipo de turmas que os filhos têm. Durante te o dia na escola, os teus pais irão a aulas semelhantes às tuas, mas não à tua actual turma. Como poucos pais conhecem o circuito da escola, os pais terão um estudante guia que os acompanhará de aula para aula, que lhes explique onde ir, lhes responda às questões e que almoce com eles.

5. Situação de uma nomeação para liderar o grupo

Na tua turma todos têm de realizar um trabalho de grupo onde é suposto estudarem história de uma forma mais divertida. Para este trabalho os grupos de 5 a 10 elementos devem escolher um tema que retrate algum acontecimento importante da História de Portugal. Cada grupo deve ter um director que organize e coordene os esforços do grupo.

6. Situação de um colega (par) conselheiro

O psicólogo da tua escola está a tentar desenvolver um novo programa em que estudantes com problemas podem procurar outros estudantes, assim como adultos, para os ajudarem. Estes estudantes chamar-se-ão pares conselheiros. O psicólogo procurou-te a ti e a outros alunos para sugerirem alguns colegas da vossa turma, para esta função. Eles disseram que procuram pessoas em que os outros estudantes sintam que podem confiar para exporem os seus problemas. Também procuram alguém que seja bom ouvinte e que se preocupe de facto com os colegas.

Na versão original, a ordem pela qual as situações aparecem é aleatória e cada prova é acompanhada de uma lista com os nomes de todos os colegas da turma. Para cada situação é pedido ao jovem que nomeie três rapazes e três raparigas da sua turma que considere mais competentes para atingir os objectivos especificados em cada situação. Depois de fazerem as nomeações, solicita-se aos sujeitos que avaliem a sua própria competência naquela situação. As classificações são expressas numa escala de cinco pontos desde o *muito bom* ao *fraco*. Os professores também usam essa escala para classificarem a eficácia comportamental dos alunos nas situações. Para cada conjunto de resultados é obtida a média nas seis situações de forma a obter um índice global de competência social. Barnes e Sternberg (1989) incluíram uma nova questão com o objectivo de recolher o índice de complexidade percebido pelo indivíduo na resolução de cada situação, oscilando desde muito complexo até ao nada complexo.

Esta prova tem sido utilizada em diferentes estudos como critério de avaliação da eficácia comportamental em situações sociais, com populações de adolescentes (Ford, 1982; Ford & Tisak, 1983) e de adultos (Barnes & Sternberg, 1989), nos EUA. No estudo de adaptação portuguesa, usámos a versão original de Ford (1982), dirigida a jovens estudantes entre o 7.º e o

12.º ano de escolaridade acrescentando algumas das sugestões propostas por Barnes e Sternberg (1989) na versão que desenvolveram para adultos. Assim, começámos por traduzir e adaptar o texto da prova para a língua portuguesa (Nota 1). De seguida, apresentámo-la a um grupo de 8 professores, 4 psicólogos, 4 investigadores e 8 jovens (4 raparigas e 4 rapazes, dos 8.º e 11.º ano de escolaridade) com o objectivo de analisar a adequação e a pertinência dos conteúdos situacionais para a população portuguesa. Também foi analisada a compreensibilidade das situações, ambiguidade, credibilidade, clareza e objectividade da redacção das situações e das questões para o objectivo do estudo (Almeida & Freire, 2000; Angleitner & Wiggins, 1986).

Mantiveram-se algumas das características das versões americanas, nomeadamente a solicitação aos indivíduos para que nomeiem três colegas (rapazes ou raparigas da turma) que considerem os mais competentes para atingir os objectivos especificados em cada situação, que avaliem a sua própria competência naquela situação (auto-avaliação de desempenho), e o grau de dificuldade que a situação oferece (auto-avaliação da complexidade da situação), numa escala de cinco pontos. Por sua vez, considerou-se necessário proceder a algumas adaptações e alterações, por exemplo usamos seis situações reformuladas e reestruturadas em termos de conteúdos; optámos por não utilizar a lista de nomes dos colegas da turma, pois considerámos que desta forma, o jovem poderia fazer escolhas em função da informação mais acessível na sua memória; os conteúdos e formato das situações-estímulo foram adaptados e reformulados em função das opiniões recolhidas e foi inserida uma situação de exemplo; não forçamos a escolha de três rapazes e de três raparigas, nem a sua ordenação (apenas a nomeação de três colegas da turma percebidos como os mais competentes para aquele papel); e utilizamos esta prova para os jovens se auto-avaliarem e aos colegas, assim como para recolher as avaliações dos professores. No quadro II transcrevemos as situações usadas na prova portuguesa.

Quadro 2 - Lista das situações incluídas na prova portuguesa**1. Situação de trabalho de casa**

Nesta escola, todos os alunos se queixam, porque este ano os professores passaram trabalhos de casa para as férias do Natal. Os alunos desta turma, mais do que queixarem-se ou ficarem zangados, juntaram-se e pediram aos professores para ouvirem o seu lado da história. Um grupo de professores concordou em conversar com um representante dos alunos sobre as suas razões na próxima reunião de professores.

2. Situação de um duplo encontro

Há um aluno(a) bastante contente porque finalmente conseguiu um encontro com alguém de quem gosta à muito tempo. Porém essa pessoa irá acompanhada pelo(a) primo(a), que durante esta semana se encontra de visita em sua casa. Esse aluno(a) vem pedir-lhe ajuda para escolher alguém da turma que o(a) acompanhe ao encontro e que na altura adequada se retire com o(a) primo(a) e o(a) deixe a sós com a pessoa de quem gosta.

3. Situação de uma nomeação para liderar um grupo

Esta turma tem de realizar trabalhos de grupo onde é suposto estudarem literatura portuguesa de uma forma mais divertida. Para este trabalho os grupos de 5 a 10 elementos devem escolher um tema que retrate algum acontecimento importante da literatura portuguesa. Cada grupo deve ter um director ou líder que organize e coordene os esforços do grupo.

4. Situação de um colega (par) conselheiro

O psicólogo e assistente social da escola estão a tentar desenvolver um novo programa em que estudantes com problemas podem procurar outros estudantes para os ajudarem. Estes estudantes chamar-se-ão pares conselheiros e devem ser pessoas em quem os outros estudantes sintam que podem confiar para exporem os seus problemas. Também procuram alguém que seja bom escutante e que se preocupe de facto com os colegas.

5. Situação de reunião de pais

Esta turma está muito entusiasmada com a possibilidade de fazer uma viagem de finalistas para um destino exótico para a qual já tem a verba necessária. Os pais de alguns alunos não concordam com o destino escolhido pela turma pois é demasiado longe. Se esses alunos não tiverem autorização dos pais, a viagem não se poderá realizar, pois eles também contribuíram para angariar o dinheiro que paga a viagem. A turma decidiu pedir a esses pais para virem à escola ouvir os argumentos da turma. Os pais concordaram em conversar com um representante da turma sobre a viagem de finalistas e o destino escolhido. A turma vem pedir-lhe ajuda para escolher um representante da turma para a reunião com os pais.

6. Situação de um estudante visitante

Um(a) dos(as) professores(as) teve um grave acidente. Os colegas da turma juntaram-se e decidiram fazer algo. A turma pede-lhe ajuda para decidir quem escolher (na turma) para fazer uma visita ao(à) professor(a) e à sua família. Essa pessoa deverá levar chocolates/flores e tentar expressar o quanto os estudantes lamentam o sucedido e desejam uma rápida recuperação.

Resultados

A análise dos resultados evidencia que a escala global apresenta uma média de 74.0 e um desvio-padrão de 17.55. Em termos gerais, a maioria dos itens revelou uma distribuição adequada ao longo dos vários pontos da escala *likert* utilizada. A análise do poder discriminativo (PD) dos itens revelou correlações do item com o total da escala corrigido (*ritc*) com valores superiores a .20 (valor considerado crítico para a dimensão da amostra de referência, $p < .01$). No Quadro III descrevemos esses coeficientes, assim como o índice de consistência da escala quando se eliminam os itens (*alpha se...*).

Quadro 3 - *Estudo da correlação do item com a escala (corrigido)*

Itens	<i>ritc</i>	Alpha se
1.1	0,288	0,872
1.2	0,387	0,869
1.3	0,450	0,867
1.4	0,527	0,866
2.1	0,350	0,870
2.2	0,316	0,871
2.3	0,409	0,869
2.4	0,430	0,868
3.1	0,344	0,871
3.2	0,392	0,869
3.3	0,417	0,868
3.4	0,499	0,866
4.1	0,437	0,868
4.2	0,491	0,866
4.3	0,542	0,864
4.4	0,624	0,864
5.1	0,359	0,870
5.2	0,363	0,870
5.3	0,435	0,868
5.4	0,558	0,865
6.1	0,453	0,867
6.2	0,469	0,867
6.3	0,511	0,865
6.4	0,632	0,863

O estudo da consistência interna, através do coeficiente alpha de Cronbach (Nunnally, 1978), revelou um valor de alfa de .87, sendo mais elevado que o valor encontrado por Ford e Tisak (1983) para uma amostra de 600 adolescentes americanos ($\alpha = .76$). Os auto-registos de competência apresentam um índice de consistência interna mais baixa comparativamente aos registos dos pares ou dos professores, o que poderá sugerir que *"self-judgments are more situation specific than judgments by others, which tend to be more traitlike"* (Ford, 1982, p.330).

As nomeações pelos pares (NP) dos colegas com melhor nível de competência ou facilidade para resolver a situação, intra-turma, realizou-se em função do número/frequência de nomeações de cada indivíduo, para o total das seis situações. No nosso estudo, optámos por transformar cada resultado individual numa nota z, indicadora da posição ocupada pelo indivíduo face aos seus pares (intra-turma), e que torna possível outras comparações. Por sua vez, as avaliações dos professores tomam as seis situações seleccionadas para o contexto português, pela ordem em que são apresentadas na versão definitiva. Os professores nomeiam os três alunos que, em seu entender, terão maior facilidade e maior dificuldade para cada uma das situações.

De seguida, procedemos ao estudo da *validade de construto*, que assenta na análise em componentes principais (ACP), usando o método de rotação *varimax*, considerando os itens da prova dirigida aos jovens. No Quadro IV apresentam-se os valores obtidos. Os testes preliminares da matriz de intercorrelação ($KMO = 0,87$) e de esfericidade de Bartlett ($\chi^2 = 3009,201$; $gl = 276$; $p < .001$) apresentam índices adequados. Seis factores apresentam valor-próprio igual ou superior à unidade, explicando no seu conjunto 64% da variância dos itens (o primeiro factor explica 19,0%).

Quadro 4 - Análise em componentes principais com rotação varimax

Itens	Componentes					
	1	2	3	4	5	6
1.1	0,148	-0,112	0,322	-0,200	0,661	0,005
1.2	-0,039	-0,041	0,351	0,113	0,701	0,045
1.3	0,634	-0,023	0,320	0,004	0,251	-0,090
1.4	0,394	0,106	0,459	0,117	0,279	0,098
2.1	0,067	0,723	0,035	0,044	0,144	0,002
2.2	-0,164	0,840	0,042	0,112	0,149	0,034
2.3	0,497	0,551	-0,002	0,034	-0,073	0,063
2.4	0,273	0,773	0,092	0,046	-0,021	0,074
3.1	0,096	-0,095	0,723	0,059	0,099	0,152
3.2	-0,056	0,100	0,761	0,260	0,026	0,110
3.3	0,636	-0,043	0,394	0,075	0,025	-0,093
3.4	0,264	0,188	0,711	0,142	0,019	0,113
4.1	0,140	-0,032	0,077	0,766	0,135	0,102
4.2	0,011	0,116	0,229	0,851	0,130	0,029
4.3	0,601	0,134	0,100	0,432	0,012	0,057
4.4	0,298	0,156	0,240	0,680	0,131	0,170
5.1	0,098	-0,062	0,073	0,085	0,129	0,817
5.2	-0,099	0,112	0,205	0,092	0,103	0,835
5.3	0,673	0,056	-0,013	0,040	-0,064	0,453
5.4	0,437	0,178	0,178	0,148	0,073	0,574
6.1	0,142	0,189	-0,169	0,255	0,673	0,152
6.2	0,005	0,308	-0,077	0,314	0,678	0,120
6.3	0,584	0,266	-0,206	0,225	0,278	0,125
6.4	0,421	0,301	-0,007	0,291	0,500	0,166
Valores próprios	6,444	2,353	1,927	1,612	1,555	1,459
% Variância explicada	26,848	9,806	8,029	6,715	6,480	6,079

Nota: As saturações sombreadas foram consideradas fundamentais para a interpretação dos factores. A rotação converge em 10 iterações.

As saturações obtidas sugerem a existência de um primeiro factor que podemos identificar como um factor geral de desempenho em situações sociais (Factor I). Os cinco factores seguintes reúnem itens associados aos conteúdos situacionais, replicando a teoria subjacente à construção da escala: o Factor II reporta-se à competência social em situações de *relacionamento íntimo* com pares; o Factor III associa-se aos indicadores de competência em situações de *comunicação interpessoal e liderança* em contexto escolar; o Factor IV remete para uma dimensão da competência social associada ao *apoio e comunica-*

ção *inter-pares*; o Factor V está associada ao *apoio a adultos*; e o Factor VI tem a ver com a *comunicação e interacção com pais*.

Estes resultados sugerem que a percepção de competência social se dirige quer a um componente geral de desempenho social quer a aspectos mais situacionais. Neste sentido, podemos assumir que os itens desta prova são semelhantes às situações sociais vivenciadas pelos sujeitos, reproduzindo a sua especificidade. Este facto está de acordo com a intenção inicial de criação de provas situacionais para avaliar a competência social, no sentido de garantir a sua validade de construto e a sua validade ecológica.

Finalmente, procedemos à correlação dos resultados com outras variáveis dos alunos (validade de critério). Os resultados na escala de *competência social percebida* pelo jovem, baseada na nota global da TCS, apresentam-se *positivamente correlacionados com a origem sócio-económica e com o nível de participação social* dos alunos ($r=.25$; $p<.001$). A correlação positiva e significativa entre a competência social percebida pelo jovem e o índice de participação social constitui um novo critério de validade da TCS, no sentido em que prognostica fenómenos comportamentais em meio natural (critério de *validade ecológica*, de acordo com Goldstein, 1996; Samuda, Feuerstein, Kaufman, Lewis, & Sternberg, 1998; Sternberg, 1998).

Cruzando esses resultados com as classificações dos alunos em várias disciplinas escolares, os coeficientes de correlação obtidos aproximam-se de .00 e em nenhum caso com significado estatístico. A correlação entre a competência social percebida pelo jovem e as suas dificuldades sociais percebidas pelos professores situou-se em $-.17$ ($p<.01$), sendo interessante quer a sua significância estatística quer o sentido inverso da correlação obtida (convergência de indicadores). Por último, a facilidade percebida pelos pares correlaciona-se em .20 com a competência social percebida pelo jovem, o que vem de reforçar, mais uma vez, a validade ecológica da prova.

Conclusões

A necessidade crescente de se ultrapassar a visão reducionista da inteligência confinada a uma leitura intelectual ou racional, justifica o interesse crescente pela operacionalização e avaliação de formas alternativas de inteligência ou, pelo menos, de comportamentos inteligentes diversos. Assim emergiram na investigação psicológica os conceitos de inteligência social e de inteligência prática, entre outros (Gardner, 1983, 1999; Greenspan & Driscoll, 1997; Sternberg & Grigorenko, 2003). Nesta altura, num misto de variáveis pessoais e contextuais, o termo “inteligência” dá muitas vezes lugar ao conceito de “competência”.

Reportando-nos à competência social, uma das teorias mais profícuas de operacionalização e instrumento para a sua avaliação prende-se com os trabalhos de Greenspan (Greenspan, 1979, 1981; Greenspan & Driscoll, 1997). Tratando-se de uma área e de instrumentos recentes, e não tendo Portugal grande tradição de construção ou adaptação de instrumentos de avaliação psicológica, consideramos oportuno avançar para a adaptação da *Social Competence Nomination Form* (SCNF – Ford, 1982), assumindo as suas três versões (avaliação pelo próprio, pelos pares e pelos professores).

Após análises qualitativas assentes na compreensão e na relevância dos itens junto de alunos, professores e técnicos, a versão adaptada para a população portuguesa foi objecto de análises estatísticas tendo em vista apreciar as qualidades psicométricas da escala, nomeadamente a consistência interna dos seus itens e a validade (validade interna, externa e ecológica). Assim, a consistência interna da escala atinge um nível adequado de coeficiente *alpha*, ao mesmo tempo que a análise da estrutura factorial dos itens sugere a existência de um primeiro factor que se confunde com a competência social geral e os restantes associados aos conteúdos ou especificidades das situações sociais em avaliação. Por outro lado, cruzando as auto-avaliações dos alunos e as avaliações feitas pelos pares e pelos professores, os valores mostram-se estatisticamente significativos, sugerindo a validade ecológica deste teste de competência social. Quando se tomam as auto-avaliações dos alunos, e as percepções de professores e pares face a critérios externos de participação social, os resultados encontrados também se apresentam estatisticamente significativos, corroborando a validade externa e a validade ecológica da prova.

A finalizar, destacamos a pertinência do conceito de competência ser assumido com um significado amplo, integrando quer os elementos operativos (funções cognitivas) quer os elementos informativos, associados ao indivíduo e ao contexto, que pressupõem uma avaliação da competência auto-percebida, da competência hetero-percebida e do desempenho social. A prova aqui apresentada, nas suas versões para alunos, professores e pares, apresenta-se como uma opção válida para diagnosticar dificuldades e potencialidades dos alunos em situações diversificadas do quotidiano social e interpessoal. Esta prova pode, por isso, constituir um dos instrumentos a integrar o processo de avaliação psicológica de jovens, com vista ao seu desenvolvimento e/ou formação.

Nota 1 – A tradução e retroversão da escala contou com o apoio de Guilhermina Rebocho, Professora de Português/Inglês no Departamento de Pedagogia e Educação da Universidade de Évora, que muito se agradece.

Referências

- Almeida, L. & Freire, T. (2000). *Metodologia da investigação em psicologia e educação* (2.^a ed.). Braga: Psiquilíbrios.
- Almeida, L. & Roazzi, A. (1988). Inteligência: A necessidade de uma definição e avaliação contextualizada. *Psychologica, 1*, 93-104.
- Angleitner, A. & Wiggins, J. (1986). *Personality assessment via questionnaires: Current issues in theory and measurement*. Berlin: Springer-Verlag.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Barnes, M. & Sternberg, R. J. (1989). Social intelligence and decoding of nonverbal cues. *Intelligence, 13*, 263-287.
- Candeias, A. A. & Almeida, L. S. (1999). Inteligência social: Contributos para uma psicologia mais contextualizada. *Mente Social, 4* (1), 121-146
- Candeias, A. A. (2001). Inteligência Social: Estudos de conceptualização e operacionalização do construto. *Tese de Doutoramento* (Documento policopiado). Évora: Universidade de Évora.
- Cantor, N. & Kihlstrom, J. (1987). *Personality and social intelligence*. Hillsdale, NJ: Prentice-Hall.
- Cantor, N. & Kihlstrom, J. (1989). Social intelligence and cognitive assessment of personality. In R. Wyer & Th. Scrull (Eds.), *Advances in social cognition* (Vol. II, pp. 1-60). Hillsdale, NJ: Lawrence Erlbaum.
- Ford, M. (1982). Social cognition and social competence in adolescence. *Developmental Psychology, 18* (3), 323-340.
- Ford, M. (1986). For all practical purposes: Criteria for defining and evaluating practical intelligence. In R. Sternberg & R. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday competence* (pp. 183-201). Cambridge: Cambridge University Press.
- Ford, M. (1995). Intelligence and personality in social behavior. In D. Saklofske & M. Zeidner (Eds.), *International handbook of personality and intelligence* (pp. 125-140). New York: Plenum Press.
- Ford, M. & Tisak, M. (1983). A further search for social intelligence. *Journal of Educational Psychology, 75* (2), 196-206.
- Gardner, H. (1983). *Frames of Mind*. New York: Basic Books.
- Gardner, H. (1999). *Intelligence reframed*. New York: Basic Books.
- Gibbs, J. & Widaman, K. (1982). *Social intelligence: Measuring the development of socio-moral reflection*. London: Prentice-Halls.

- Goldstein, G. (1996). Functional considerations in neuropsychologist. In R. Sbordonne & C. Long (Eds.), *Ecological validity of neuropsychological testing* (pp. 75-90). Florida: GR Press/St. Lucie Press.
- Greenspan, S. (1979). Social intelligence in the retarded. In N. R. Ellis (Ed.), *Handbook of mental deficiency, psychological theory and research* (2^a ed.) (pp. 483-531). Hillsdale, NJ: Lawrence Erlbaum.
- Greenspan, S. (1981). Defining childhood social competence: A proposed working model. In B. Keogh (Ed.), *Advances in special education* (Vol. III, pp. 1-82). Greenwich, CT: JAI Press.
- Greenspan, S. & Driscoll, J. (1997). The role of intelligence in a broad model of personal competence. In D. Flanagan, J. Genshaft & P. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 131-150). London: The Guilford Press.
- Gresham, F. & Elliot, S. (1990). *Social skills rating system*. Circle Pines, MN: American Guidance Service.
- Guilford, J. P. (1959). Three faces of intellect. *American Psychologist*, 14, 469- 479.
- Guilford, J. P. (1985). The structure of intellect model. In B. Wolman (Ed.), *Handbook of intelligence* (225-266). New York: John Wiley & Sons.
- Kihlstrom, J. & Cantor, N. (2000). Social Intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 359-379). Cambridge: Cambridge University Press.
- Mathias, J. & Nettelbeck, T. (1992). Reliability of seven measures of social intelligence in a sample of adolescents with mental retardation. *Research in Developmental Disabilities*, 13, 131-143.
- Nunnally, J. C. (1978). *Psychometric theory* (2nd. ed.). New York: McGraw-Hill.
- Oppenheimer, L. (1989). The nature of social action: Social competence versus social conformism. In B. Schneider (Ed.), *Social competence in developmental perspective* (Vol. 51, pp. 41-70). NATO ASI Series.
- Pelechano, V. (1994). Assessment of interpersonal skills across the life-span. *Revue Européenne de Psychologie Appliquée*, 44 (4), 281-287
- Samuda, R. J., Feuerstein, R. Kaufman, A. S., Lewis, J. E. & R. J. Sternberg (1998). *Advances in cross-cultural assessment*. Thousand Oaks, CA: Sage Publications.
- Selman, R. (1976). Toward a structural analysis of developing interpersonal relations concepts: Research with normal and disturbed preadolescents boys. In A. D. Dick (Ed.), *Minnesota Symposia on Child Psychology* (Vol. 10). Minneapolis: University of Minnesota Press.
- Selman, R. (1980). *The growth of interpersonal understanding: Clinical and developmental analyses*. New York: Academic Press.
- Shure, M. & Spivak, G. (1988). Interpersonal cognitive problem solving. In R. Price, E. Cohen, R. Lorin & J. Mackay (Eds.), *Fourteen ounces of prevention: A case book for practitioners* (pp. 69-82). Washington: APA.

- Spivack, G., Platt, J. & Shure, M. (1976). *The problem-solving approach to adjustment*. San Francisco: Jossey-Bass.
- Sternberg, R. J. (1998). All intelligence testing is “cross-cultural.” In R. J. Samuda e Cols. (Eds.), *Advances in cross-cultural assessment* (pp. 197-215; 274-285). Thousand Oaks, CA: Sage Publications.
- Sternberg, R. J. & Grigorenko, E. (2003). *Evaluación dinámica del potencial de aprendizaje*. Barcelona: Paidós.
- Stricker, L. & Rock, D. (1990). Interpersonal competence, social intelligence, and general ability. *Personal and Individual Differences*, 11 (8), 833-839.
- Turiel, E. (1980). The development of social-conventional and moral concepts. M. Windmiller, N. Lambert & E. Turiel (Eds.), *Moral development and socialization* (pp. 69-106). Boston: Allyn & Bacon.
- Turiel, E. (1983). Domains and categories in social-cognitive development. In W. Overton (Ed.), *The relationship between social and cognitive development* (pp. 53-90). Hillsdale, NJ: Lawrence Erlbaum.
- Yeates, K. O. & Selman, R. L. (1989). Social competence in the schools: Toward an integrative developmental model for intervention. *Developmental Review*, 9, 64-100.

SOCIAL COMPETENCE: ITS EVALUATION ON DEVELOPMENTAL AND EDUCATIONAL CONTEXTS

Adelinda Araújo Candeias

Universidade de Évora

Leandro S. Almeida

Universidade do Minho

Abstract: Social competence, as a domain of practical intelligence, currently presents more importance on aptitudes and cognitive functioning analysis. Its definition includes cognitive and contextual aspects, which must be considered in psychological tests for its assessment. This paper presents a Portuguese social competence tests based on self perceptions, and on peers and teachers nominations to assess the young social competence in order to solve specific social problems. A sample of 326 students of 7th to 11th grades was used to assess the items' internal consistency and factor components. Results showed a general social competence factor, as well as five other factors related to the specific situations or items contents (intimacy, leadership, relationship with peers, adults and parents). Zero correlation coefficients were obtained between social competence and school marks, as well as positive and significant correlation coefficients with personal social competence and peers and teachers nominations.

KEY-WORDS: *Social competence; intelligence; social problem-solving; cognitive assessment.*

AVALIAÇÃO DA CARREIRA EM CONTEXTO ESCOLAR

José Manuel Tomás da Silva

Faculdade de Psicologia e de Ciências da Educação, Universidade de Coimbra

Resumo

Neste estudo fazemos uma revisão do estado actual da avaliação de carreira em Portugal. A avaliação de carreira é definida como o processo através do qual se procura obter informação sobre as características das pessoas e dos seus contextos de vida com o objectivo de as ajudar a conhecerem-se melhor e a resolverem os seus problemas de escolha e de ajustamento à carreira. Nesse sentido ao longo deste trabalho são sumariamente apresentados alguns dos principais instrumentos usados na avaliação de carreira com a finalidade de auxiliar na previsão, discriminação e monitorização do comportamento e do desenvolvimento vocacional.

PALAVRAS-CHAVE: *Avaliação de Carreira, Inteligência, Aptidões, Interesses, Valores, Maturidade, Personalidade.*

Introdução

A avaliação da carreira é, tradicionalmente, um dos componentes mais importantes da intervenção de carreira (e.g., Betz, 1992; Chartrand e Walsh, 2001; Hackett e Watkins, 1995; Hartung, 2005; Prediger, 1974; Savickas, 2000; Spokane, 1991). De facto, pelo menos desde que Frank Parsons, no início do século passado, lançou as bases modernas da Orientação Vocacional, considerando que a escolha avisada de uma profissão requeria: (a) o aprofundamento do conhecimento do self, (b) o conhecimento da estrutura de oportunidades profissional, e (c) a aplicação dos princípios e dos métodos de tomada de decisão para ponderar adequadamente os dois tipos de conhecimentos, que os profissionais da orientação têm procurado desenvolver modelos, métodos e técnicas de avaliação psicológica, para auxiliarem os seus clientes a conhecerem melhor as características pessoais relevantes para a realização e a adaptação às escolhas vocacionais. Esse empreendimento, com a sua forte ênfase na avaliação, sobretudo, a de tipo psicotécnica, marcou profundamente a evolução da Psicologia e da Orientação Vocacional durante

Morada (address): José Manuel Tomás da Silva. Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra, Rua do Colégio Novo, 3001-802 Coimbra, Portugal. e-mail: jtsilva@fpce.uc.pt

a primeira metade do século vinte, facto unanimemente reconhecido pelos autores que se têm dedicado à análise histórica dos primórdios desta disciplina. Estes autores consideram que o “movimento dos testes”, cuja origem se pode encontrar nos estudos de Galton, Binet e Cattell, e o desenvolvimento da Orientação Vocacional estão indissolivelmente interligados. Hoje podemos afirmar, com confiança, que na primeira metade do século vinte a Psicologia Vocacional foi, essencialmente, uma Psicologia das Diferenças individuais aplicada ao contexto vocacional ou da carreira (Dawis, 1992; Savickas e Barker, 2005). Ora, como se sabe, o ramo da psicologia, que estuda as diferenças individuais no comportamento, caracteriza-se por uma extensiva utilização dos métodos e técnicas de avaliação psicológica, donde a enorme influência que os testes e outras modalidades de observação psicológica tiveram na psicologia vocacional. Embora no início da segunda metade do século passado, o modelo *Traço-e-Factor*, enquanto expressão no domínio vocacional da psicologia das diferenças individuais, tenha começado a perder alguma ascendência, que outrora detinha no seio das abordagens do comportamento vocacional, cedendo parte da sua influência a novas abordagens de cariz mais fenomenológico e neo-comportamentalista, esta teoria e modelo de organização das práticas de orientação, actualmente reformulada como teoria do ajustamento pessoa-ambiente (e.g., Dawis, 2005; Holland, 1997), continua a ter um papel marcante na psicologia vocacional, quer a nível da investigação, quer no plano dos serviços de carreira. No entanto, durante os anos sessenta e setenta houve, indiscutivelmente, um nítido abrandamento na utilização dos procedimentos clássicos de avaliação psicométrica. A psicologia vocacional acompanhava, assim, a tendência que nessa altura perpassou praticamente todos os campos da psicologia, sob a influência de eminentes detractores da teoria dos traços como, por exemplo, Mischel (1968). O cepticismo e negativismo reinante sobre o valor do conceito de traço transformou-se rapidamente numa crítica incisiva aos testes psicológicos, dada a estreita associação destes com a investigação sobre os traços de personalidade. Neste clima de desconfiança sobre os testes psicológicos e acerca do seu valor para a resolução dos problemas dos indivíduos, muitos profissionais de orientação deixaram, nessa altura, de utilizar instrumentos de avaliação psicométrica, enquanto outros, passaram a usá-los mais esporadicamente. Alguns conselheiros de carreira, em virtude deste facto, começaram a recorrer com maior frequência a métodos e a técnicas de avaliação qualitativa e informal. Algumas dessas técnicas, como a entrevista, já eram frequentemente usadas na orientação vocacional, mas muitas outras (e.g., desenho de linhas da vida, genogramas de carreira, *card sorts*), constituíram interessantes novidades, no domínio da

avaliação de carreira e, desde então, têm recebido uma boa aceitação por parte de alguns profissionais de orientação. Todavia, estes meios “não tradicionais” de avaliação, nunca se constituíram como real alternativa dos procedimentos mais clássicos, de cariz mais quantitativo (e.g., testes, questionários e inventários, escalas de auto-avaliação). A passagem do tempo mostra, indubitavelmente, que o anúncio da morte, tanto da abordagem traço-e-fator, como da abordagem quantitativa da avaliação psicológica na orientação e no aconselhamento de carreira (e.g., Goldman, 1961) foi uma premonição precipitada e claramente exagerada, atendendo à popularidade de que os testes e outros procedimentos de medida psicológicos, desfrutaram na actualidade, tanto entre os profissionais como no seio dos clientes dos serviços de carreira. De facto, quer uns, quer outros, apoiam a sua utilização e reconhecem o interesse da informação produzida para aumentar o auto-conhecimento bem como para a realização de decisões, em diferentes áreas da vida das pessoas. Como salientaram Chartrand e Walsh (2001), a “avaliação de carreira sempre foi a parte mais importante do aconselhamento psicológico, mas nunca foi tão largamente aceite como é presentemente” (p. 231). Outro indicador da vitalidade do campo da avaliação de carreira pode ser encontrado no aumento dos meios inteiramente dedicados à divulgação dos instrumentos e da investigação realizada neste domínio. Neste âmbito refira-se, como exemplo paradigmático, o aparecimento em 1993 do *Journal of Career Assessment*, uma revista científica cabalmente dedicada à difusão de trabalhos na esfera da avaliação de carreira. Na mesma linha de ideias, a monografia *A Counselor's Guide to Career Assessment Instruments* (Kapes e Mastie, 1982), publicada pela primeira vez no começo dos anos oitenta e cuja edição tem vindo a ser regularmente actualizada, dadas as rápidas mudanças que continuam a ocorrer, a nível do desenvolvimento de instrumentos de avaliação de carreira e nas profissões que os utilizam, representa uma das principais fontes de informação existentes, no fórum internacional, sobre os instrumentos de avaliação de carreira.

No nosso país também encontramos exemplos como aqueles que acabamos de referir para os EUA. As monografias *Provas Psicológicas em Portugal* (Almeida, Simões e Gonçalves, 1995), *Testes e Provas Psicológicas em Portugal* (Simões, Gonçalves e Almeida, 1999), e, mais recentemente, os dois volumes da Colecção Avaliação Psicológica, da Quarteto Editora (Almeida, Simões, Machado e Gonçalves, 2004; Gonçalves, Simões, Almeida e Machado, 2003), embora sendo genericamente dedicados à avaliação psicológica, contêm, apesar disso, informação útil e relevante acerca de instrumentos utilizados na avaliação de carreira, devendo por isso ser consideradas obras de re-

ferência neste domínio. O livro *Avaliação Psicológica em Orientação Escolar e Profissional* (Leitão, 2004), também publicado recentemente pela mesma editora atrás mencionada, e reunindo os contributos de vários autores, nacionais e internacionais, especializados na área da avaliação psicológica, constitui a fonte mais actualizada e completa de que dispomos no nosso país, até à data, sobre a avaliação de carreira.

De acordo com um trabalho de revisão recente sobre os instrumentos de avaliação utilizados, em Portugal, na investigação e na intervenção de carreira (Teixeira, 2004a), pode-se comprovar que o seu número e diversidade, nomeadamente, no que diz respeito às características psicológicas e de carreira que são avaliadas, é bastante significativo e abrangente. Sem ter tido a preocupação de fazer um levantamento exaustivo do conjunto de instrumentos de carreira existentes no contexto nacional, mesmo assim, a autora, reuniu e apresentou uma lista composta por quarenta instrumentos distintos que, no agregado, permitem medir as dimensões tradicionalmente consideradas relevantes nas intervenções de carreira: testes de inteligência e de aptidões cognitivas específicas, inventários de interesses e de valores acerca do trabalho, medidas do desenvolvimento/maturidade de carreira e avaliações da personalidade. Noutro estudo, mais antigo, que também procedeu a um levantamento dos instrumentos de avaliação psicológica mais frequentemente utilizados pelos psicólogos dos *Serviços de Psicologia e de Orientação* (SPO's), pertencentes à *Direcção Regional de Educação do Centro* (DREC), os autores (Cussecala e colaboradores, 1999), identificaram uma extensa lista de vinte e oito instrumentos que, permitem medir um largo espectro de características individuais.

No espaço limitado de que dispomos para este trabalho não seria viável descrever cada um dos instrumentos identificados nos estudos, anteriormente referidos, pelo menos de um modo minimamente razoável e sério. Assim, neste texto iremos efectuar uma revisão selectiva de um pequeno conjunto de instrumentos, comumente utilizados na avaliação da carreira em contexto escolar, em Portugal. Para a escolha dos instrumentos utilizámos os dois seguintes critérios: (1) as provas deverão estar validadas para a população portuguesa e, cumulativamente, sempre que possível, que possam ser obtidas facilmente pelos interessados (i.e., que estejam disponíveis no mercado), e (2) constem em pelo menos uma das fontes anteriormente citadas.

Antes, porém, de avançarmos para a descrição dos instrumentos seleccionados precisamos de esclarecer, ainda que sumariamente, o que entendemos por avaliação de carreira e as suas principais funções a nível das intervenções de carreira, nomeadamente, as que se realizam em contexto escolar.

Definição e Principais Funções da Avaliação de Carreira

A avaliação de carreira é um domínio especializado da avaliação psicológica. Logicamente, entre os dois domínios existem dilatadas áreas de sobreposição, tanto a nível dos conteúdos, métodos e técnicas, como no plano dos princípios, competências e padrões éticos e deontológicos. A maioria dos psicólogos vocacionais estaria de acordo em definir avaliação psicológica como um processo que envolve a recolha de informação com a finalidade de ajudar as pessoas a compreenderem e a defrontarem os seus problemas (Walsh e Betz, 2001). Nesse sentido, pode dizer-se que avaliação de carreira constitui um processo central na intervenção de carreira “proporcionando aos psicólogos dados objectivos acerca dos clientes e dos seus problemas” (Savickas, 2004), no domínio específica da carreira ou, mais geralmente, do papel do trabalho na vida das pessoas.

Os psicólogos vocacionais realizam avaliações de carreira, com duas intenções principais, segundo Spokane (1991). A primeira está em correspondência com o objectivo psicométrico tradicional de testar e classificar as pessoas, com a finalidade de prever futuros comportamentos de interesse, de testar teorias formais do comportamento vocacional, ou de estudar a epidemiologia de problemas sociais ou psicológicos (objectivo “actuarial”). A segunda, representando uma meta geral mais familiar da avaliação na intervenção de carreira, consiste em ajudar os clientes a realizarem decisões de carreira (objectivo de utilidade aplicada). Assim, de acordo com o autor citado, a avaliação de carreira, é habitualmente executada com o propósito de: (1) pôr a descoberto possibilidades de carreira congruentes com a personalidade vocacional do indivíduo, (2) determinar a linha dominante dos interesses, para confirmar a congruência de preferências ou de escolhas, (3) realizar diagnósticos na avaliação de conflitos ou de problemas, (4) motivar comportamentos construtivos de carreira, (5) que os clientes adquiriram uma estrutura cognitiva que os ajude a valorarem as alternativas de carreira, (6) utilizar informação no esclarecimento de expectativas e na planificação de tratamentos e, por último, (7) determinar a amplitude (ou espectro) de aptidão do indivíduo.

Para concretizarem estes diferentes propósitos, os psicólogos vocacionais utilizam uma diversificada gama de procedimentos de avaliação psicológica, habitualmente incluídos nas quatro amplas categorias, a seguir indicadas: Testes de Inteligência e de Aptidões Específicas, Inventários de Interesses e de Valores, Medidas de Desenvolvimento/Maturidade de carreira e Testes de Personalidade.

Um psicólogo vocacional só raramente precisará de utilizar instrumentos extraídos das quatro categorias referidas. Na maioria das vezes optará por usar uma selecção restrita de procedimentos de medida, englobando instrumentos dirigidos para duas ou três características da personalidade, com relevância para a decisão e adaptação à carreira. Lowman e Carson (2000), por exemplo, propõem um modelo “inter-domínios” de avaliação e aconselhamento, que privilegia a apreciação integrada de interesses-aptidões-personalidade (cf., também, com os trabalhos de Ackerman e Heggestad, 1997). Super (1983), por sua vez, defendeu as virtudes de se efectuar uma avaliação de *precisão* (por oposição à metodologia dita de *saturação*) dos problemas, das pessoas e dos seus contextos de vida, no modelo *Desenvolvimentista de Avaliação e Aconselhamento de Carreira* (C-DAC). Segundo este modelo holístico de avaliação e aconselhamento, o psicólogo começaria por efectuar um levantamento do problema (ou objectivo) do cliente, recorrendo a medidas do desenvolvimento/maturidade de carreira, e utilizaria esses resultados preliminares para determinar a necessidade de uma avaliação em profundidade da personalidade do indivíduo (por exemplo, dos interesses, aptidões, valores, e outros traços de personalidade). Face ao que acabámos de dizer, parece-nos essencial que o psicólogo vocacional, detenha um sólido conhecimento dos principais instrumentos incluídos em cada uma das quatro categorias anteriormente identificadas, e que demonstre proficiência na sua utilização e aplicação, no contexto da intervenção de carreira. De seguida e, para concluir este trabalho, iremos apresentar alguns instrumentos, validados para a população portuguesa, que avaliam distintas características humanas, relevantes para a compreensão e explicação do papel do trabalho na vida das pessoas e para o desenvolvimento e a intervenção de carreira.

Testes de Inteligência e de Aptidões Específicas (Multifactoriais)

A avaliação da inteligência e de outras aptidões cognitivas constitui, praticamente, desde a origem da Psicologia científica, nos finais do século dezanove, um domínio privilegiado da atenção dos psicólogos. Embora, tanto as provas de inteligência, como as das aptidões específicas (nomeadamente, as integradas em baterias multifactoriais), avaliem aspectos do funcionamento cognitivo da personalidade, os dois conceitos, e os instrumentos que lhes estão associados, aparecem habitualmente distinguidos na literatura especializada (Almeida, 1988; Almeida e Simões, 2004; Cohen e Swerdlik, 2002; Walsh e

Betz, 2001). Esquemáticamente, poderia dizer-se que enquanto o conceito de inteligência remete para designações como “inteligência geral, ou factor *g*”, “capacidade intelectual”, ou o popularizado “quociente de inteligência, ou *QI*”, pretendendo traduzir o funcionamento unitário e integrado de distintas, mas inter-relacionadas, componentes cognitivas, já o constructo de “aptidão” tende a ser definido de um modo claramente diferenciado. De facto, segundo Freeman (1962/1980), uma aptidão “é a combinação de características que apontam a capacidade de um indivíduo para adquirir (mediante treino) um conhecimento específico, uma perícia, um conjunto de respostas organizadas, como a aptidão para falar uma língua, para ser músico, ou para realizar determinadas tarefas mecânicas. Assim, um teste de aptidão visa medir a *capacidade potencial de um indivíduo numa actividade especializada*, dentro de um âmbito restrito” (p. 463).

Ambos os instrumentos são utilizados na avaliação de carreira, como auxiliares na previsão e no prognóstico diferencial do sucesso do cliente nas diferentes trajectórias de vida/carreira, por si contempladas. Dentro deste objectivo comum, os primeiros serão, sobretudo, utilizados para determinar a probabilidade de sucesso na prossecução de estudos superiores e pós-graduados e em profissões mais exigentes, em termos educativos e, concomitantemente, requerendo um maior grau de autonomia, complexidade e responsabilidade da parte dos seus ocupantes. Enquanto que os segundos vão ser utilizados, primordialmente, para prognosticar a futura performance do indivíduo em distintas áreas educativas, em cursos de formação profissional e em empregos específicos, em função da avaliação de distintas aptidões (apreciadas singularmente ou integradas em baterias multifactoriais). Este objectivo corresponde, aliás, ao processo tradicional, na orientação vocacional, de emparelhamento das características da pessoa com as exigências do emprego.

Os psicólogos vocacionais portugueses dispõem de uma ampla gama de provas, devidamente validadas em grupos nacionais, para realizarem uma avaliação das capacidades cognitivas dos seus clientes e para efectuarem previsões (probabilistas) acerca do seu eventual grau de sucesso em distintos percursos educativos e profissionais. Atendendo ao facto de, neste trabalho, apenas abordarmos os processos de avaliação em contexto escolar, fundamentalmente, dirigidos para jovens do 9º ao 12º anos de escolaridade, apenas referiremos dois testes, que se encontram entre os mais habitualmente usados para estes fins.

A *Bateria de Provas de Raciocínio Diferencial* (BPRD) é, de acordo com os dados de que dispomos, uma das provas do funcionamento cognitivo mais utilizadas na avaliação de carreira, em Portugal. A BPRD surge na linha dos tra-

balhos de investigação iniciados por Leandro S. Almeida nos anos oitenta, com os Testes de Raciocínio Diferencial (TRD), de Meuris. Esta bateria integra cinco provas, cada uma delas pretendendo avaliar “a capacidade de raciocínio (operação cognitiva comum) num determinado conteúdo específico dos seus itens (numérico, abstracto, verbal, espacial e mecânico)” (Almeida, 1995, p. 19). As provas foram validadas junto de alunos entre o 7º e 12º ano de escolaridade. As análises psicométricas dos principais parâmetros estatísticos dos resultados têm revelado, em sucessivos estudos, que as cinco escalas de raciocínio (NR-Numérico, AR-Abstracto, VR-Verbal, SR-Espacial e MR-Mecânico) apresentam bons níveis, quer de fidelidade (coeficientes de consistência interna, no intervalo de .80-.90), quer de validade de constructo, concorrente e preditiva. Particularmente interessante para o uso deste instrumento na orientação vocacional, são as correlações encontradas entre os distintos testes e as classificações escolares em várias disciplinas (Matemática, Português, Físico-Química, etc.). Por exemplo, nos alunos do 10º ao 12º ano, o autor (e.g., Almeida, 1988) constatou que as correlações mais elevadas se verificaram entre as provas e as disciplinas com conteúdos similares (e.g., SR e Geometria Descritiva). Existem normas (nomeadamente, sobre a forma de “notas T”, “notas percentílicas” e “classes normalizadas”), elaboradas segundo o ano de escolaridade, o sexo e o meio urbano/rural, aos quais se acrescentam as “opções escolares”, para os alunos do ensino secundário. O tempo para a administração da prova completa é de aproximadamente 60 minutos. A BPRD é, actualmente, distribuída pela CEGOC – Investigação e Publicações Psicológicas.

A *Bateria de Testes de Aptidões (GATB)* é a versão Portuguesa da *General Aptitude Test Battery*, um instrumento de aptidão multifactorial desenvolvido pelo *Serviço de Emprego dos Estados Unidos (USES)*, para ser utilizada nos vários serviços de emprego estatais. Para alguns autores (e.g., Whiston, 1999), a GATB merece ser discutida e conhecida, porque se trata do teste de aptidão ocupacional mais extensivamente investigado. A introdução deste teste em Portugal deve-se ao laborioso e profícuo trabalho de investigação realizado por Helena R. Pinto, da Universidade de Lisboa (e.g., Pinto, 1999, 2004). Existem algumas diferenças entre a GATB original e a adaptação Portuguesa (de seguida, apenas referida de Bateria), esta última reunindo somente oito dos 12 testes de papel e lápis que integram a versão norte-americana. A Bateria é composta pelos testes seguintes: Comparação de nomes; Cálculo numérico; Desenvolvimento de volumes; Vocabulário; Utensílios idênticos; Raciocínio aritmético; Emparelhar formas; Fazer três traços. As características metrológicas da Bateria foram profundamente investigadas numa grande amostra de aferição que incluiu 12.970 estudantes, do 9º ao 12º anos de es-

colaridade. De especial interesse são os resultados sobre a estrutura factorial dos testes da Bateria. Nos vários estudos realizados, através da análise factorial, foram extraídos três factores que explicam, aproximadamente, dois terços da variância total. Os três factores avaliam, respectivamente, aspectos Simbólicos, Perceptivos e Burocrático-Motores (aliás, estes correspondem aos três resultados compósitos de grupos de aptidões que podem ser derivados a partir dos resultados nos 8 testes da Bateria). Quanto ao estudo da precisão, a autora refere que “os coeficientes *alpha* de Cronbach revelaram índices elevados de consistência interna das provas (entre .64 e .88)” (Pinto, 1999, p. 65). Também no âmbito dos estudos diferenciais, realizados com os testes da Bateria, pode comprovar-se que o instrumento permite diferenciar “favoravelmente os níveis de escolaridade mais avançados, os estudantes com notas escolares mais elevadas, cujos pais desempenham profissões científicas e técnicas e que frequentam escolas de zonas urbanas; diferenciam também vias e áreas de formação; e, relativamente a algumas aptidões, diferenciam os dois sexos” (Pinto, 1999, p. 66). Com base no estudo de aferição foram produzidas tabelas de normas para as oito aptidões e para os grupos de aptidões, sendo possível calcular-se resultados padronizados de média 100 e desvio padrão 20, para o conjunto das (quatro) subamostras de standardização e para cada sexo. A bateria é composta por um numeroso conjunto de testes de aptidões, cuja aplicação completa demora aproximadamente 1,30 horas. A edição e distribuição nacional da GATB está a cargo da autora.

Para além dos instrumentos analisados anteriormente, muitos outros estão validados na nossa população, podendo por isso ser usados com interesse na avaliação de carreira, de seguida referimos apenas mais dois exemplos:

- PMA – *Aptidões Mentais Primárias* (L. L. Thurstone; CEGOC-TEA)
- WISC-III – Escala de Inteligência de Wechsler para Crianças, 3ª edição (D. Wechsler; CEGOC-TEA)

Inventários de Interesses e de Valores Vocacionais

A avaliação dos interesses e dos valores relativos ao trabalho e à carreira constitui um dos principais contributos da Psicologia Vocacional para o domínio da avaliação psicológica. É um facto bem conhecido que a maior parte da investigação sobre a medida dos interesses tem sido realizada por psicólogos vocacionais, desde que nos finais dos anos vinte, do século passado, E. K. Strong, Jr. desenvolveu o inventário de interesses mais provecto da história da avaliação de carreira. Os interesses e os valores (juntamente com

as necessidades) são características do funcionamento psicológico, geralmente, classificadas no domínio da motivação humana. A motivação refere-se aos processos e factores que permitem explicar aspectos importantes do comportamento, como a direcção, o esforço e a persistência ou perseverança perante situações desfavoráveis, ou adversas. Por vezes os conceitos de interesse e de valor são tratados como equivalentes, todavia, constituem constructos distintos (Dawis, 1991; Rokeach, 1973). Os *valores* são os objectivos que as pessoas almejam alcançar, enquanto que os *interesses* são as actividades e os objectos através das quais os indivíduos procuram atingir os objectivos, ou metas, quer no trabalho/carreira, quer noutras esferas da vida. Dito de outra forma, os valores revelam-nos o que é importante para as pessoas (e.g., independência, variedade, prestígio), ao passo que os interesses revelam os meios que os indivíduos gostariam de utilizar para alcançarem os seus objectivos finais (e.g., ser o Administrador Geral de uma das 10 principais empresas nacionais, para alguém que valoriza o prestígio). Tradicionalmente, a avaliação dos interesses tem predominado sobre o exame dos valores. Todos os anos milhares de estudantes portugueses respondem a inventários de interesses vocacionais, com o objectivo de “descobrirem” quais os seus gostos ou preferências, a respeito de múltiplas actividades profissionais. Em contrapartida, os inventários de valores só esporadicamente serão avaliados nos SPO’s, na maioria das escolas nacionais. No entanto, com as transformações e mudanças aceleradas nas carreiras, determinadas pelo advento da sociedade pós-industrial, alguns autores defendem que a avaliação dos valores, rapidamente alcançará uma posição central, no conjunto das estratégias que serão mobilizadas pelos psicólogos para ajudarem os seus clientes a gerirem eficazmente as suas carreiras.

A avaliação dos interesses vocacionais, no nosso país, está menos desenvolvida que a área da inteligência e das aptidões, apesar dos esforços apreciáveis que têm sido feitos por alguns autores, desde a década de noventa, do século passado, (Ferreira e Hood, 1995; Leitão e Miguel, 2001; Taveira, Nunes, Mesquita, Alves e Milhazes, 1995; Teixeira, 2000). A avaliação neste domínio, crucial para uma escolha vocacional cientificamente fundada, é, ainda neste momento, realizada em muitos SPO’s, com base em versões experimentais, não comercializadas, de inventários de interesses. De seguida, vamos analisar, com maior detalhe, dois dos mais populares instrumentos desta categoria.

California Occupational Preference System – Interest Inventory (COPS) [renomeado *Career Occupational Preference System*, ou *COPSystem Interest Inventory*: Knapp e Knapp, 1976; Knapp-Lee, 2000]. O Inventário de Interesses

do COPSsystem faz parte de um programa compreensivo de orientação de carreira que, para além dos interesses vocacionais, avalia, igualmente, as capacidades (CAPS: *Career Ability Placement Survey*) e os valores (COPES: *Career Orientation Placement and Evaluation Survey*), tendo sido desenvolvido com a intenção de ajudar os indivíduos (alunos do ensino pré-universitário, universitário e adultos) no processo de tomada de decisão de carreira. Em Portugal, existe apenas uma versão experimental do Inventário de Interesses (COPS) que, apesar do seu estatuto algo provisório, tem sido amplamente utilizada na avaliação de carreira de jovens e jovens adultos (segundo os dados do inquérito realizado por Cussecala e colaboradores, acima referido, o COPS é o inventário de interesses mais frequentemente utilizado pelos Psicólogos integrados na CAE de Coimbra). Actualmente, a EdITS comercializa diversas formas, revistas e novas, para avaliação dos interesses (cf. Knapp-Lee, 2000). O COPS foi construído, nos finais dos anos 60, como medida dos interesses dos indivíduos nas oito dimensões, ou *clusters* de interesses, propostas por A. Roe: Ciência (A); Tecnologia (B); Economia (D); Secretariado (E), Arte (F); Serviços (G); Ar Livre (J); Comunicação (L). Para além destes oito agrupamentos principais existem outros seis sub-*clusters*, que resultam do desdobramento de alguns dos grupos anteriores, e dizem respeito a actividades profissionais com níveis educacionais mais baixos de preparação: Ciência (Preparação técnica, H); Tecnologia (Preparação Técnica, I; Economia de mercado e de consumo diário, C); Economia (Preparação técnica, K); Arte (Preparação técnica, M) e Serviços (Preparação técnica, N).

Taveira e colaboradores (1995) realizaram o primeiro estudo de validação deste instrumento numa amostra de 852 alunos de ambos os sexos, que frequentavam o 9º, 10º, 11º e 12º anos de escolaridade em estabelecimentos de ensino, públicos e privados, de Braga. Recorrendo à análise factorial, os autores extraíram 9 factores que contribuíram para explicar cerca de 52.1% da variância total. Com a excepção do factor 9 (integrando maioritariamente itens do *cluster* Economia de mercado e consumo diário), a maior parte dos restantes itens mostraram correlações significativas num dos oito agrupamentos principais (a sub-divisão por níveis de formação não foi corroborada). Os autores concluíram que a consistência interna dos nove *clusters* era suficiente, justificando a utilização do instrumento com fins aplicados (alfas entre .83 e .94).

Interesses e preferências Profissionais (IPP: De La Cruz, 1993/2001). Este inventário, comercializado em Portugal pela CEGOC-TEA, foi desenvolvido em Espanha, no início dos anos 90, por De La Cruz, e desde que foi adaptado para Portugal tem vindo a registar uma forte utilização na avaliação de

carreira, por parte dos profissionais dos SPO's (e.g., Cussecala *et. al.*, 1999). O inventário permite avaliar os interesses dos indivíduos, jovens (a partir dos 13 anos) e adultos, em 17 campos profissionais (e.g., Científico-Experimental, Científico-Técnico, Literário, etc.), utilizando como estímulos títulos de profissões e actividades/tarefas profissionais. O relatório técnico apresenta informação detalhada sobre o estudo de adaptação realizado no nosso país, nomeadamente, estatísticas relativas a diversos parâmetros dos itens e das escalas do inventário (por exemplo, os coeficientes alfa de Cronbach, variam entre .61 e .85, encontrando-se a maioria acima de .70). O manual português contém tabelas com normas diferenciadas, para os distintos campos, segundo o ano de escolaridade (do 9º ao 12º ano) e o género. Existe uma disquete para a realização da cotação automática das respostas.

Há outros instrumentos validados para a população portuguesa que, mesmo não sendo tão utilizados como os dois referidos nesta súmula, merecem maior atenção no futuro, por parte dos Psicólogos Vocacionais. A título de exemplo, destacamos os seguintes:

- *Inventário de Interesses Vocacionais de Jackson* (Teixeira, 2004b)
- *Inventário de Interesses Vocacionais* (Ferreira e Hood, 1995)
- *Inventário de Interesses Vocacionais de Amsterdam* (Evers, Miguel e Leitão, 2000).

As medidas dos valores do Trabalho/Vida podem ser utilizadas na avaliação de carreira, como os interesses, para ajudar os indivíduos a discriminarem entre diferentes fileiras educativas (e.g., áreas de estudos, cursos) e distintos sectores de actividade profissional, defendendo alguns autores, inclusive, que os valores profissionais suplantarão os interesses nessa função. Apesar de serem menos frequentemente avaliados que os interesses, há, em Portugal, uma forte linha de investigação sobre os valores, que se tem desenvolvido no contexto de estreitas relações internacionais (e.g., Duarte, 2004; Marques, 1980/81/82, 1995). A construção de escalas de avaliação dos valores e de outros instrumentos similares (e.g., o *Inventário sobre a Saliência das Actividades*) foram alguns dos resultados importantes do envolvimento dos portugueses no *Work Importance Study* (WIS).

A *Escala de Valores WIS* (edição de 1981) permite avaliar 21 valores (e.g., Utilização das Capacidades, Realização, Promoção, Estético), nove dos quais podem considerar-se intrínsecos e os restantes extrínsecos. Dentro de cada escala respectiva, os primeiros cinco itens revestem-se de carácter geral e os cinco últimos estão estreitamente ligados ao trabalho. O estudo de validação nacional incidiu sobre 3 amostras (estudantes do 9º, do 12º ano e adul-

tos), tendo as análises revelado que a precisão das escalas é, em geral, adequada (as medianas dos coeficientes alfa nas três amostras são .83, .85 e .79). Outros estudos realizados, nomeadamente, no âmbito de dissertações de mestrado, têm mostrado a robustez teórica e empírica dos constructos psicológicos avaliados.

Para além da *Escala de Valores WIS* existem outras escalas e inventários validados para a população portuguesa, nomeadamente:

- *Inventário de Valores Pessoais* (L. V. Gordon; CEGOC-TEA)
- *Inventário de Valores de Vida* (LVI: Almeida e Fraga, 2004)

Medidas do Desenvolvimento/Maturidade de Carreira

As medidas analisadas até agora permitem recolher dados com base nos quais os indivíduos podem realizar *predições* do sucesso em comportamentos educativos e profissionais (testes de inteligência e de aptidões) e efectuar *discriminações*, isto é, permitem que os indivíduos descubram quais os grupos escolares e profissionais com que mais se assemelham (inventários de interesses e de valores). Em conjunto, as funções preditiva e discriminativa dos testes usam informação pertinente para o *conteúdo* das escolhas vocacionais (Betz, 1992; Herr et al., 2004). Existe, todavia, uma outra classe de medidas que lida com informação relevante para o *processo* de escolha vocacional, isto é, o tipo de informação que permite avaliar como os sujeitos estão a abordar o processo de escolha, qual o seu estágio de desenvolvimento nesse processo e as dificuldades com que se debatem. Neste caso os instrumentos de avaliação são usados com uma finalidade de *monitorização* (Herr et al., 2004). Para monitorizarem os comportamentos de carreira, de indivíduos e de grupos, os psicólogos dispõem de uma enorme variedade de procedimentos de medida, recaindo os mais comuns nas categorias de instrumentos de avaliação das dificuldades de escolha (Silva, 2004a) e da maturidade de carreira (Silva, 2004b).

O *Inventário dos Factores de Carreira* (IFC: e.g., Silva, 1997) constitui uma adaptação do *Career Factors Inventory* (CFI), desenvolvido nos EUA, nos finais da década de oitenta, por J. Chartrand, S. Robbins, W. Morrill e K. Boggs. Trata-se de uma escala multidimensional da indecisão de carreira, relativamente curta (composta por apenas 21 itens), e que pode ser usada na intervenção de carreira, para examinar o contributo de quatro factores para a dificuldade do sujeito em realizar uma escolha: Ansiedade na Escolha; Indecisão Generalizada; Necessidade de Informação sobre a Carreira; e Necessidade de Conhecimento do Self. Os estudos realizados em Portugal corrobora-

ram a qualidade metrológica do instrumento, quer no que respeita à precisão, quer quanto a distintos aspectos de validade (Lima, 1998; Silva, 1998).

O *Inventário do Desenvolvimento de Carreira* (IDC: Marques e Caeiro, 1979), é a adaptação portuguesa do *Career Development Inventory* (CDI) de Super, Thompson, Lindeman, Jordaan e Myers, um instrumento que pretende avaliar a prontidão (maturidade/adaptabilidade) do sujeito – a forma escolar, Forma S, destina-se aos alunos do 9º ao 12º ano de escolaridade – para enfrentar as tarefas de desenvolvimento de carreira próprias do estágio de exploração vocacional. O inventário é composto por quatro escalas que avaliam os seguintes elementos da maturidade de carreira: Planeamento de Carreira, Exploração de Carreira, Tomada de Decisão e Informação sobre o Mundo do Trabalho (Há ainda uma quinta escala – Conhecimento da Profissão Preferida – que só raramente é utilizada, dadas as dificuldades que engendra para a construção de resultados normativos). Existe uma abundante informação psicométrica acerca do IDC, proveniente, principalmente, de estudos realizados em dissertações de mestrado.

Estão disponíveis, para uso em Portugal, muitos instrumentos do processo de escolha de carreira, apenas a título de exemplo, referimos os seguintes:

- *Escala de Exploração Vocacional* (Taveira, 1997)
- *Escala de Auto-Eficácia Ocupacional* (Paixão e Silva, 2004)
- *Inventário de Crenças de Carreira* (Rafael, 2001)
- *Inventário de Percepção de Barreiras de Carreira* (Cardoso, 1999).

Testes de Personalidade

Os instrumentos de personalidade durante muitas décadas não foram utilizados com regularidade na avaliação de carreira. A baixa frequência no uso destas medidas, provavelmente, justifica-se pela dúvida, lançada por investigações realizadas entre os anos 50 e 80, sobre a aplicação dos testes de personalidade em vários contextos da intervenção de carreira. Só recentemente, parece ter começado a emergir uma tendência mais favorável acerca da utilização destes instrumentos de avaliação na psicologia vocacional (e.g., Walsh e Eggerth, 2005). Um dos factores que parece mais ter contribuído para este interesse foi a emergência e a popularidade do modelo dos “Cinco Factores”, como esquema organizativo da personalidade. Embora as variáveis de personalidade tenham sido usadas para prever a performance em tarefas profissionais, os psicólogos vocacionais, como referem Zunker e Osborne (2002), recorrem aos inventários de personalidade “para medir dife-

renças individuais em traços sociais, impulsos motivacionais e necessidades, atitudes e grau de ajustamento – informação vital no processo de exploração vocacional” (p. 113).

O Inventário de Personalidade NEO Revisto (NEO PI-R) é um dos instrumentos disponíveis no mercado português (editado e distribuído pela CEGOC-TEA), que apresenta maiores virtualidades para a utilização na avaliação de carreira. O NEO PI-R trata-se da adaptação nacional do inventário homónimo desenvolvido nos estados unidos por Paul T. Costa e Robert R. McCrae. O inventário tem sido alvo de um extenso processo de validação em amostras nacionais, sobretudo, seleccionadas da população adulta (e.g., Lima e Simões, 1995), e permite efectuar uma profunda avaliação de 5 domínios (Neuroticismo, Extroversão, Abertura à Experiência, Amabilidade e Conscienciosidade) e de 30 facetas da personalidade (Lima, 1997). A investigação tem revelado que as escalas incluídas no inventário possuem precisão e validade suficientes para avaliarem aspectos relevantes do desenvolvimento de carreira (nomeadamente, os relativos à satisfação e ao bem estar, em diferentes domínios profissionais).

Os psicólogos dispõem ainda de uma ampla gama de instrumentos de personalidade para utilizarem na avaliação de carreira, entre os quais destacamos os seguintes:

- *CEP, Questionário de Personalidade* (J. L. Pinilos; CEGOC-TEA)
- *Inventário de Personalidade PRF-A* (Teixeira, 2000).

Conclusão

A avaliação de carreira é uma das áreas mais dinâmicas e desenvolvidas da Psicologia Vocacional. Os psicólogos vocacionais fazem avaliações de carreira para múltiplos fins, e porque consideram que a informação obtida pode revelar-se extremamente útil para ajudar as pessoas de diferentes idades, raças, género e estatuto sócio-económico, a realizarem decisões de carreira que tenham uma boa probabilidade de reverter em sucesso, satisfação e bem estar para o próprio e, indirectamente, para a comunidade e sociedade em que estas se inserem. Para efectuarem avaliações de carreira, os psicólogos recorrem a uma enorme gama de procedimentos objectivos e subjectivos, formais e informais, quantitativos e qualitativos, embora, geralmente, os testes e outros instrumentos quantitativos estejam entre os mais frequentemente utilizados. Neste trabalho analisámos e apresentámos, de forma necessariamente resumida, alguns dos principais instrumentos utilizados em quatro categorias distin-

tas: inteligência e aptidões; interesses e valores; desenvolvimento e maturidade de carreira; personalidade.

Em Portugal, como pensamos ter demonstrado ao longo deste trabalho, existe actualmente uma oferta considerável de instrumentos para a avaliação de carreira, contudo, consideramos ser necessário continuar a investigar e a produzir nesta área, por forma a dotarmos os profissionais de orientação de meios, inovadores e actualizados, que lhes permitam intervir mais eficazmente no processo de escolha e de adaptação vocacional dos portugueses.

Referências

- Ackerman, P. L. e Heggstad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin*, 121, 219-245.
- Almeida, L. S. (1988). *O raciocínio diferencial dos jovens*. Porto: INIC.
- Almeida, L. S. (1995). Bateria de Provas de Raciocínio Diferencial (BPRD). In L. S. Almeida, M. R. Simões e M. M. Gonçalves (Eds.), *Provas Psicológicas em Portugal* (pp. 19-28). Braga: APPORT.
- Almeida, L. S. e Simões, M. R. (2004). Os testes de inteligência na orientação vocacional. In M. L. Leitão (Coord.), *Avaliação psicológica em orientação escolar e profissional* (pp. 79-100). Coimbra: Quarteto.
- Almeida, L. S., Simões, M. R. e Gonçalves, M. M. (Eds.) (1995). *Provas psicológicas em Portugal*. Braga: APPORT.
- Almeida, L. S., Simões, M. R., Machado, C. e Gonçalves, M. M. (Coords.) (2004). *Avaliação psicológica instrumentos validados para a população portuguesa. Vol. II*. Coimbra: Quarteto.
- Almeida, M. L. e Fraga, S. (2004). Os valores em estudantes do 12º ano de escolaridade: Estudos com o Life Values Inventory (LVI). In M. C. Taveira, H. Coelho, H. Oliveira e J. Leonardo, *Desenvolvimento vocacional ao longo da vida. Fundamentos, princípios e orientações* (pp. 269-276). Coimbra: Almedina.
- Betz, N. E. (1992). Career assessment: A review of critical issues. In S. D. Brown e R. W. Lent (Eds.), *Handbook of counseling psychology* (2ª ed., pp. 453-484). New York: Wiley.
- Cardoso, P. (1999). *Percepção de barreiras na carreira e atitudes face à exploração e planeamento de carreira em alunos do ensino secundário*. Dissertação de mestrado. Lisboa: Universidade de Lisboa.
- Chartrand, J. M., e Walsh, W. B. (2001). Career assessment: Changes and trends. In F. T. L. Leong e A. Barak (Eds.), *Contemporary models in vocational psychology. A volume in honor of Samuel H. Osipow* (pp. 231-255). Mahwah, NJ: Lawrence Erlbaum Associates.
- Cohen, R. J. e Swerdlik, M. E. (2002). *Psychological testing and assessment. An introduction to testing and assessment* (5ª ed.). Boston: McGraw-Hill.
- Cussecala, A., Sampaio, A., Sequeira, F., Vieira, G., Cabrita, M., Konigstedt, M., Varandas, P. e Felício, V. (1999). *Avaliação psicológica. Um estudo sobre a prática nas escolas*. Coimbra: DREC – Direção Regional de Educação do Centro.
- Dawis, R. V. (1991). Vocational interests, values, and preferences. In M. D. Dunnette e L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (vol. 2, pp. 833-871). Palo Alto, CA: Consulting Psychologists Press.
- Dawis, R. V. (1992). The individual differences tradition in counseling psychology. *Journal of Counseling Psychology*, 39, 7-19.

- Dawis, R. V. (2005). The Minnesota theory of work adjustment. In S. D. Brown e R. W. Lent (Eds.), *Career development and counseling. Putting theory and research to work* (pp. 3-23). Hoboken, NJ: John Wiley & Sons, Inc.
- De La Cruz, M. V. (2001). *IPP. Interesses e Preferências Profissionais. Manual* (A. F. Barros, trad., 3ª ed.). Lisboa: CEGOC-TEA (obra original publicada em 1993).
- Duarte, M. E. (2004). A avaliação dos valores. In M. L. Leitão (Coord.), *Avaliação psicológica em orientação escolar e profissional* (pp. 155-178). Coimbra: Quarteto.
- Evers, A., Miguel, J. P. e Leitão, M. L. (2000). Un inventario para la evaluación de los intereses en el nivel universitario: Estudio de adaptación del Inventario de Intereses Vocacionales de Amsterdam (IIVA). In Gabinet d'Orientació Universitària, *Congrés d'Orientació Universitària* (pp. 215-229). Barcelona: Edicions Universitat de Barcelona.
- Ferreira, J. A. e Hood, A. B. (1995). The development and validation of a Holland-Type Portuguese vocational interest inventory. *Journal of Vocational Behavior*, 46, 119-130.
- Freeman, F. S. (1980). *Teoria e prática dos testes psicológicos* (M. J. Miranda, trad.). Porto: Fundação Calouste Gulbenkian (Obra original publicada em 1962).
- Goldman, L. (1961). *Using tests in counseling*. New York: Appleton-Century-Crofts.
- Gonçalves, M. M., Simões, M. R., Almeida, L. S. e Machado, C. (Coords.) (2003). *Avaliação psicológica instrumentos validados para a população portuguesa. Vol. I*. Coimbra: Quarteto.
- Hackett, G. e Watkins, C. E. (1995). Research in career assessment: Abilities, interests, decision making, and career development. In W. B. Walsh e S. E. Osipow (Eds.), *Handbook of vocational psychology: Theory, research, and practice* (2ª ed., pp. 181-215). Mahwah, NJ: Lawrence Erlbaum Associates.
- Hartung, P. J. (2005). Integrated career assessment and counseling: Mindsets, models, and methods. In W. B. Walsh e M. L. Savickas (Eds.), *Handbook of vocational psychology: Theory, research, and practice* (3ª ed., pp. 371-395). Mahwah, NJ: Lawrence Erlbaum Associates.
- Herr, E. L., Cramer, S. H. e Niles, S. G. (2004). *Career guidance and counseling through the lifespan: Systematic approaches* (6ª ed.). Boston: Pearson Education, Inc.
- Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3ª Ed.). Lutz, FL: Psychological Assessment Resources, Inc.
- Kapes, J. T. e Mastie, M. M. (1982). *A counselor's guide to vocational guidance instruments*. Falls Church, VA: National Vocational Guidance Association.
- Knapp, R. R. e Knapp, L. (1976). *COPsystem Technical Manual*. San Diego, CA: EdITS.
- Knapp-Lee, L. (2000). A complete career guidance program: The COPSystem. In C. E. Watkins, Jr. e V. L. Campbell (Eds.), *Testing and assessment in counseling practice* (2ª Edição, pp. 295-338). Mahwah, NJ: Lawrence Erlbaum Associates.

- Leitão, M. L. (Coord.) (2004). *Avaliação psicológica em orientação escolar e profissional*. Coimbra: Quarteto.
- Leitão, L. M. & Miguel, J. P. (2001). Os interesses revisitados, *Psychologica*, 26, 79–104.
- Lima, M. P. (1997). NEO-PI-R: *Contextos teóricos e psicométricos – “Ocean” ou “iceberg”?* Dissertação de doutoramento não publicada. Coimbra: Universidade de Coimbra.
- Lima, M. P. e Simões, A. (1995). Inventário de Personalidade NEO PI-R. In L. S. Almeida, M. R. Simões e M. M. Gonçalves (Eds.), *Provas Psicológicas em Portugal* (pp. 133-149). Braga: APPORT.
- Lima, M. R. (1998). *Orientação e desenvolvimento da carreira em estudantes universitários. Estudo das atitudes de planeamento e exploração, identidade vocacional, saliência dos papéis e factores de carreira*. Dissertação de doutoramento. Lisboa: Universidade de Lisboa.
- Lowman, R. L. e Carson, A. D. (2000). Integrating assessment data into career counseling. In D. A. Luzzo (Ed.), *Career counseling for college students. An empirical guide to strategies that work* (pp.121-136). Washington, DC: APA.
- Marques, J. F. (1980/81/82). A investigação psicológica sobre valores. Desenvolvimento de uma nova escala de valores. *Revista Portuguesa de Psicologia*, 17/18/19, 135-158.
- Marques, J. F. (1995). The Portuguese Work Importance Study. In D. E. Super e B. Sverko (Eds.), *Life roles, values and careers. International findings of the Work Importance Study*. (pp. 181-187). San Francisco: Jossey-Bass Publishers.
- Marques, J. H. e Caeiro, L. A. (1979). O Inventário de Desenvolvimento Vocacional (CDI) em Portugal: Estudo Preliminar. *Biblos*, LV, 501-514.
- Mischel, W. (1968). *Personality and assessment*. New York: Willey.
- Paixão, M. P. e Silva, J. M. T. (2004, Abril). *Psychometric characteristics of the Portuguese version of the Occupational Self-Efficacy Scale*. Comunicação (poster) apresentada no VII Congreso Europeo de Evaluación Psicológica, Málaga.
- Pinto, H. R. (1999). Bateria de Testes de Aptidões GATB. In M. R. Simões, M. Gonçalves e L. S. Almeida (Eds.), *Testes e provas psicológicas em Portugal* (Vol. II; pp. 60-70). Braga: SHO/APPORT.
- Pinto, H. R. (2004). Avaliação das aptidões e orientação vocacional. In L. M. Leitão (Coord.), *Avaliação psicológica em orientação escolar e profissional* (pp. 123-151). Coimbra: Quarteto.
- Prediger, D. J. (1974). The role of assessment in career guidance. In E. L. Herr (Ed.), *Vocational guidance and human development* (pp. 325-349). Washington, DC: Houghton Mifflin Company.

- Rafael, M. (2001). *O modelo desenvolvimentista de avaliação e aconselhamento da carreira* (C-DAC). Dissertação de doutoramento não publicada. Lisboa: Universidade de Lisboa.
- Rokeach, M. (1973). *The nature of human values*. New York: Free Press.
- Savickas, M. L. (2000). Assessing career decision making. In C. E. Watkins, Jr., e V. L. Campbell (Eds.), *Testing and assessment in counselling practice* (2ª ed., pp. 429-480). Mahwah, NJ: Lawrence Erlbaum Associates.
- Savickas, M. L. (2004). Um modelo para a avaliação de carreira. In M. L. Leitão (Coord.), *Avaliação psicológica em orientação escolar e profissional* (pp. 21-42). Coimbra: Quarteto.
- Savickas, M. L. e Baker, D. B. (2005). The history of vocational psychology: Antecedents, origin, and early development. In W. B. Walsh e M. L. Savickas (Eds.), *Handbook of vocational psychology: Theory, research, and practice* (3ª ed., pp.15-50). Mahwah, NJ: Lawrence Erlbaum Associates.
- Silva, J. M.T. (1997). *Dimensões da indecisão da carreira: Investigação com adolescentes*. Dissertação de doutoramento não publicada. Coimbra: Universidade de Coimbra.
- Silva, J. M. T. (1998). Avaliação multidimensional do constructo da indecisão de carreira: O Inventário dos Factores de Carreira (IFC). *Psychologica*, 20, 41-57.
- Silva, J. M. T. (2004a). Avaliação da indecisão vocacional. In M. L. Leitão (Coord.), *Avaliação psicológica em orientação escolar e profissional* (pp. 347-386). Coimbra: Quarteto.
- Silva, J. M. T. (2004b). Avaliação da maturidade de carreira. In M. L. Leitão (Coord.), *Avaliação psicológica em orientação escolar e profissional* (pp.265-316). Coimbra: Quarteto.
- Simões, M. R., Gonçalves, M. M. e Almeida, L. S. (Eds.) (1999). *Testes e provas psicológicas em Portugal. Vol. II*. Braga: S.H. O. – Sistemas Humanos e Organizacionais, Lda.
- Spokane, A. R. (1991). *Career intervention*. Englewood Cliffs, NJ: Prentice Hall.
- Super, D. E. (1983). Assessment in career guidance: Toward truly development counselling. *Personnel and Guidance Journal*, 61, 555-562.
- Taveira, M. C. (1997). *Exploração e desenvolvimento vocacional de jovens. Estudo sobre as relações entre a exploração, a identidade, e a indecisão vocacional*. Dissertação de doutoramento não publicada. Braga: Universidade do Minho.
- Taveira, M. C., Nunes, J. M., Mesquita, A., Moreno, M. L., Silva, M. C. e Milhares, G. (1994). Avaliação das necessidades de educação vocacional de jovens: Implicações para o desenvolvimento de programas de exploração vocacional. In L. S. Almeida e I. S. Ribeiro (Eds.), *Avaliação psicológica: Formas e contextos* (pp. 395-410). Braga: Apport.

- Teixeira, M. O. (2000). *Personalidade e motivação no desenvolvimento vocacional. As necessidades, os valores, os interesses e as auto-percepções no conhecimento de si vocacional*. Dissertação de doutoramento. Lisboa: Universidade de Lisboa.
- Teixeira, M. O. (2004a). Avaliação psicológica no domínio do aconselhamento vocacional: Princípios, padrões éticos e competências. In M. C. Taveira, M. H. Coelho, H. Oliveira e J. Leonardo (Eds.), *Desenvolvimento vocacional ao longo da vida: Fundamentos, princípios e orientações* (pp. 37-56). Coimbra: Almedina.
- Teixeira, M. O. (2004b). Inventário de Interesses Vocacionais de Jackson (JVIS). In L. S. Almeida, M. R. Simões, C. Machado e M. M. Gonçalves (Coords.), *Avaliação psicológica instrumentos validados para a população portuguesa*. (Vol. II; pp. 53-68). Coimbra: Quarteto.
- Walsh, W. B. e Betz, N. E. (2001). *Tests and assessment* (4ª ed.). Upper Saddle River, NJ: Prentice Hall.
- Walsh, W. B. e Eggerth, D. E. (2005). Vocational psychology and personality: The relationship of the five-factor model to job performance and job satisfaction. In W. B. Walsh e M. L. Savickas (Eds.), *Handbook of vocational psychology* (3ª ed., pp. 267-295). Mahwah, NJ: Lawrence Erlbaum Associates.
- Whiston, S. C. (1999). *Principles and applications of assessment in counseling*. Belmont, CA: Wadsworth/Thomson Learning.
- Zunker, V. G. e Osborne, D. S. (2002). *Using assessment results for career development* (6ª ed.). Pacific Grove, CA: Brooks /Cole.

CAREER ASSESSMENT IN SCHOOL CONTEXT

José Manuel Tomás da Silva

Faculdade de Psicologia e de Ciências da Educação, Universidade de Coimbra

Abstract: The main objective of this study is to selectively review the state of the art of the career assessment field in Portugal. Career assessment is defined as the process by which we try to obtain information about the person's characteristics and the contexts in which they live, with the general goal of helping them to know themselves better and to resolve their career choice and adjustment problems. In this article we'll present, succinctly, some of the main instruments used in career assessment endeavours with the aim of prediction, discrimination and monitoring the career behavior and development of individuals.

KEY-WORDS: *Career Assessment, Intelligence, Aptitudes, Interests, Values, (Career) Maturity, Personality.*

ESTUDO NORMATIVO DO TESTE DE RORSCHACH NA POPULAÇÃO ADULTA PORTUGUESA: ASPECTOS METODOLÓGICOS E ALGUNS RESULTADOS

António Abel Pires

Faculdade de Psicologia e Ciências da Educação, Universidade do Porto

Resumo

Esta investigação apresenta alguns resultados do Estudo Normativo de Rorschach, na população adulta portuguesa, com o SIR – Sistema Integrativo de Rorschach.

A análise dos resultados da nossa amostra permite distinguir a orientação intercultural e intracultural de muitas das variáveis Rorschach. Podemos identificar variáveis do Rorschach que são sensíveis à dimensão intercultural e os seus resultados são estáveis nos três níveis de escolaridade, mas diferentes dos dados de Exner nos E.U.A. (Exner, 1993). Existem outras variáveis que são sensíveis à vertente intracultural, com diferenças significativas nos níveis educacionais. De referir que o aumento do nível educacional é responsável pelo desaparecimento de algumas diferenças entre os resultados da nossa amostra total e o Estudo Normativo dos E.U.A.

PALAVRAS-CHAVE: *Teste de Rorschach, Estudo normativo, Adultos, Portugal.*

Introdução

Os estudos normativos do Teste de Rorschach em Portugal, tanto para crianças como para adultos está em fase de conclusão. O estudo normativo para a população infantil, está a ser realizado pelo Professor Doutor Danilo Silva, da Faculdade de Psicologia e Ciências da Educação da Universidade de Lisboa, (Silva, 1991, Silva e cols., 1990, 1991, 1992, 1995, 1996, 1999, 2000, 2001, 2004). Para a população adulta, o estudo normativo tem sido realizado pelo autor e os resultados parciais e vários aspectos metodológicos foram apresentados em Congressos Internacionais de Rorschach e Métodos Projectivos (Pires, 1995, 1996, 1999, 2002, 2004, 2005).

Deve-se referir que em vários países estão a ser realizados estudos normativos, dos quais apenas indicamos: o de Espanha, região de Madrid, por Sédin (1993), Campo, na região de Barcelona (1999); Finlândia por Mattlar

e cols. (1995), Venezuela, na região de Caracas, por Ephraim, Riquelme e Occupati (1996), Chile por Vinet e cols. (1995), Bélgica por Mormont (1998), França e países francófonos por Sultan, Andronikof, Mormont e cols. (2004), na Califórnia – E.U.A. por Shaffer, Erdberg, e Haroian (1999), a repetição do estudo normativo nos E.U.A. por Exner (2002).

Neste artigo, iremos apresentar alguns aspectos metodológicos da realização do estudo normativo e discutir os resultados da amostra total e dos grupos de nível educacional.

Metodologia

Constituição da amostra

Para calcular uma amostra representativa da população adulta portuguesa tivemos em consideração as variáveis de idade, sexo, região de residência e níveis de escolaridade. A idade adulta foi considerada a partir dos 18 até mais de 75 anos. Dividiu-se o país em seis regiões: Interior norte, Litoral norte (incluindo o Grande Porto), Centro, Grande Lisboa, Interior sul e Litoral sul. As três regiões do litoral têm uma predominância de zonas urbanas pois é aí que se concentra a maioria da população e onde existem os principais centros populacionais. Nas três regiões do interior existe uma predominância de áreas rurais com pequenas zonas urbanas, correspondendo a pequenas cidades.

Para definir o nível sociocultural e socioprofissional, a classificação das profissões dava origem a uma grande quantidade de grupos. Por isso, optámos pela escolha do nível de educação formal. Esta escolha teve em consideração dois aspectos: (i) para a maioria da população, o nível de escolaridade formal permite o acesso a uma formação complementar a nível profissional ou a uma função profissional de acordo com o nível de escolaridade. Nota-se que o nível socioprofissional aumenta de acordo com o nível de escolaridade, exceptuando os casos de empresários “self made man”; (ii) no teste de Rorschach, o protocolo é constituído maioritariamente por uma produção verbal, elaborada a partir das manchas de tinta onde o sujeito mostra a sua capacidade de comunicação verbal, ao produzir as suas respostas. Poder-se-á levantar a hipótese que na maioria dos sujeitos, a capacidade de elaboração e expressão verbal, a comunicação e o à-vontade na situação de aplicação do teste poderá ser influenciada pelo nível de escolaridade e/ou de formação profissional ulterior.

Tendo em conta estes aspectos enunciados, optou-se por uma divisão em três níveis de escolaridade: o básico, correspondendo à escolaridade obriga-

tória de 0 a 6 ou de 0 a 9 anos de acordo com os grupos etários; o nível secundário, que termina ao fim de 12 anos de escolaridade e o superior, com 13 ou mais anos de escolaridade, correspondendo à frequência ou obtenção de graus universitários.

Inicialmente levantou-se a questão da patologia. O estudo normativo não pretende fixar normas rígidas, mas apenas obter resultados que permitam a percepção dos resultados reais e não necessariamente ideais de uma determinada população. Pode ser aceitável ter na amostra um determinado número de indivíduos com patologias e até se poderá respeitar na amostra a proporção de acordo com os dados estatísticos conhecidos. Isto obrigaria à constituição de uma amostra muito numerosa para se conseguir ter subgrupos representativos de cada uma das patologias.

A outra opção é obter os resultados através de uma amostra representativa de indivíduos ditos normais ou caracterizados como não pacientes e realizar um estudo posterior com grupos de diferentes patologias suficientemente representativos. A vantagem desta opção seria a de se poder obter resultados representativos dos indivíduos com diferentes patologias e poder comparar os resultados entre os diferentes grupos e com os da amostra de indivíduos ditos normais.

Para controlar a variável patologia versus normalidade optou-se pelo critério de o sujeito não ter seguido nenhum tipo de intervenção psiquiátrica ou psicológica nos últimos cinco anos e que ao mesmo tempo tenha estabilidade em termos pessoais, familiares e profissionais.

A amostra é constituída por adultos não pacientes e foi estratificada segundo o sexo, o grupo etário, o nível de escolaridade e a região de acordo com o Recenseamento da População de 1991. (INE, 1993).

Procurou-se que houvesse um número igual de homens e de mulheres apesar de no Recenseamento o número de mulheres ser ligeiramente superior (homens = 48,2%, mulheres = 51,8%). Foram escolhidos seis grupos etários dos 18 a + de 65 anos. A distribuição da amostra por idades não corresponde exactamente à da população estando os grupos mais idosos menos representados pois foi nestes que foi eliminado maior número de protocolos (ver quadro 1). A distribuição da população portuguesa por grupos etários segundo os dados do Recenseamento de 1991 é a seguinte: 18-24 = 14,9%, 25-34 = 19,3%, 35-44 = 17,6%, 45-54 = 15,3%, 55-64 = 14,8%, 65-74 = 11,0% e 75+ = 7,2%.

A escolaridade foi dividida em três níveis: básico (0-6 anos), secundário (7-12 anos) e superior (+12 anos). O nível básico corresponde à escolaridade mínima obrigatória que para os maiores de 30 anos de idade é 6

anos e para os de idade inferior a 30 anos é de 9 anos. Assim no grupo etário inferior a 30 anos é considerado o nível básico de 0-9 anos de escolaridade.

O recenseamento de 1991 indica que a maioria da população adulta tem um nível de escolaridade de 6 anos (0-6 anos = 73%, 7-12 anos = 18,5% e + 12 = 8,5%) excepto no grupo etário dos 18-25 anos (0-6 anos = 47,2%, 7-12 anos = 39,8% e + 12 anos = 13,1%) em que a maior parte deste grupo possui mais de 8 anos de escolaridade. Neste grupo etário tem-se verificado anualmente, uma progressão dos níveis de escolaridade. Os dados de 2001 (INE, 2002) indicam, para o grupo etário 18-25 anos, valores de 0-9 anos = 36%, 10-12 anos = 36% e +12 anos = 28%. Verifica-se que a população feminina se tem tornado cada vez mais numerosa na frequência do ensino superior sendo hoje maioritária em praticamente todos os cursos.

Tendo em conta a representatividade mínima de cada um destes grupos, decidimos que o nível básico representaria 50% da amostra, o nível secundário 30% e o superior 20%. Os protocolos válidos recolhidos pertencem ao nível básico = 44%, secundário = 36% e superior = 20%.

Foram escolhidas seis regiões tendo em conta as grandes regiões norte, centro e sul. Em cada uma destas grandes regiões seleccionou-se uma sub-região litoral e outra interior. Em todas as regiões tivemos em consideração a população que vive nas áreas urbanas e nas áreas rurais.

Como a grande maioria da população portuguesa é de raça branca, existindo apenas duas pequenas minorias de imigrantes negros e de ciganos, optámos nesta fase por escolher apenas sujeitos de raça branca com excepção de um sujeito mestiço (filho de mãe negra e pai branco, nascido em Angola). Para os dois grupos étnicos minoritários estamos a planificar amostras representativas de cada uma delas cujos protocolos serão recolhidos mais tarde.

O cálculo inicial da amostra previa 501 sujeitos. Foram eliminados os protocolos não válidos ou que suscitavam dúvidas na forma como o psicólogo realizou o contacto, protocolos de sujeitos que estavam a fazer qualquer tipo de tratamento médico. Retivemos 309 protocolos que foram correctamente recolhidos e considerados válidos, de acordo com as exigências deste estudo.

A redução do número de protocolos foi realizada de forma a manter as proporções da amostra planificada inicialmente.

A média de idade para a totalidade da amostra é de 37.25 (DP = 13.29, mediana = 35, moda = 26), com uma amplitude de 18 a 76 anos. Os sujeitos têm uma média de 9.33 anos de escolaridade (amplitude de 1 a 18 anos). As outras variáveis demográficas são apresentadas no quadro 1.

Quadro 1 - Variáveis Demográficas para 309 Adultos Não Pacientes

ESTATUTO MARITAL		IDADE		RAÇA	
Solteiro	84 27%	18-25 . . .	74 24%	Branca . . .	308 100%
Vive c/alg	2 1%	26-35 . . .	81 26%	Negra	0 0%
Casado	201 65%	36-45 . . .	71 23%	Asiática	0 0%
Separado	6 2%	46-55 . . .	46 15%	Outra	1 0%
Divorciado	10 3%	56-65 . . .	28 9%		
Viúvo	6 2%	> 65	9 3%		
SEXO		EDUCAÇÃO			
Masculino	155 50%	MENOS DE 7 . . .	102 33%		
Feminino	154 50%	7-9 Anos	69 22%		
		10-12 Anos	77 25%		
		+ 13 Anos	61 20%		

Procedimento

Os protocolos foram recolhidos por 7 psicólogos. Mais de 60% dos protocolos foram administrados pelo autor do presente trabalho, devido à dificuldade em encontrar profissionais correctamente formados e com disponibilidade suficiente para realizar esta tarefa.

Os protocolos foram aplicados e cotados segundo o Sistema Integrativo de Rorschach (Exner, 1993, 2001, 2003).

Depois de termos a amostra calculada, desenvolveu-se uma reflexão sobre a estratégia para contactar e seleccionar os sujeitos ao acaso e qual o local mais indicado para a aplicação do teste, tendo em conta o contexto cultural do país, as diferentes regiões e até mesmo a localidade de residência.

Foram realizadas algumas experiências e com a ajuda de psicólogos a trabalhar nos diferentes contextos, elaborou-se uma estratégia para recolher os dados.

As experiências demonstraram que o critério de selecção dos sujeitos ao acaso, através de escolha telefónica, de rua ou de números pares e ímpares não funcionava com o teste de Rorschach no contexto cultural português. Os indivíduos mostravam grande desconfiança, a maior parte recusava a colaboração, colocavam muitas questões sobre o interesse de um tal estudo, a finalidade dos resultados e a razão da sua escolha como sujeitos de investigação.

Foi afastada a hipótese do pagamento ao sujeito pela sua colaboração uma vez que a maioria recusava o pagamento e alguns consideravam-no quase um insulto.

Foi também afastada a possibilidade de apelo ao voluntariado através de anúncio de jornais, porque não nos traria os sujeitos que realmente necessitávamos mas tipos particulares de sujeitos e alguns até com alguma patologia.

Para realizar o contacto dos sujeitos elaborou-se uma estratégia de utilização de mediadores. Os psicólogos a trabalhar naquela região explicavam o objectivo deste estudo a pessoas bem inseridas na comunidade e até com algum nível de liderança, geralmente, professores, autarcas, membros de associações culturais, de forma a obter a sua cooperação para eles indicarem e apresentarem algumas pessoas podendo respeitar os critérios da amostra. O objectivo desta estratégia era a de atenuar ou eliminar a desconfiança e ao ser apresentado por um membro da comunidade em quem depositavam confiança, o contacto era mais fácil e o nível de cooperação mais elevado. Para não haver concentrações de tipos de sujeitos e respeitar na medida do possível o critério de escolha ao acaso, procurou-se aumentar o número de pessoas que poderiam indicar outros em diferentes contextos de modo a que se seleccionasse o máximo de 4 sujeitos indicados pela mesma pessoa.

No primeiro contacto com o sujeito, antes da aplicação do teste, deixava-se bem claro a liberdade de aceitar ou recusar participar no estudo e o sigilo profissional relativamente aos resultados.

A aplicação do teste foi realizada na localidade de residência do sujeito, num gabinete cedido por algum tipo de instituição e em alguns casos especiais de pessoas idosas que viviam em locais isolados, a aplicação concretizou-se na residência do sujeito.

No fim da aplicação era entregue ao sujeito uma carta de agradecimento em papel timbrado da Faculdade de Psicologia e Ciências da Educação da Universidade do Porto e assinada pelo investigador.

A maior parte dos sujeitos pediu que lhe fossem transmitidos os seus próprios resultados. Esta comunicação constitui uma forma de agradecimento pela cooperação e disponibilização do seu tempo. Os resultados foram comunicados a cada sujeito nas seguintes condições: (i) a comunicação dos resultados só era realizada após a recolha total de protocolos na localidade ou pequena região, para evitar a possível "contaminação" através de conversas e comentários sobre os resultados; (ii) como os psicólogos que aplicavam o teste não viviam na região e para evitar possíveis prejuízos provocados por resultados escritos, optou-se pela comunicação verbal dos resultados, dando a possibilidade de responder a questões e a dúvidas naquele momento. Os resultados foram comunicados de forma adequada às possibilidades de compreensão de cada sujeito.

Os dados foram tratados utilizando o RIAP 3.0 – *Rorschach Interpretation Assistance Program*, para a codificação, cálculo do sumário estrutural e exportação de ficheiros para tratamento estatístico. As estatísticas descritivas fo-

ram calculadas através do SPSS 3.0 para DOS. Os testes de normalidade (*Kolmogorov-Smirnov*), foram calculados utilizando o SPSS 8.0 para Windows. Este programa estatístico também foi utilizado para calcular as diferenças significativas entre os resultados dos diferentes subgrupos. Os testes de normalidade indicaram que a maior parte das variáveis não tinha uma distribuição normal. Em consequência foram utilizados testes não paramétricos: *Kruskal-Wallis Test* e o *Mann-Whitney Test*.

Resultados

Amostra total

A totalidade dos protocolos foi codificada de acordo com o SIR – Sistema Integrativo de Rorschach.

O acordo intercotadores é de 86%, sendo a percentagem máxima de 98% nas localizações e a mínima de 72% nas codificações especiais.

As estatísticas descritivas da nossa amostra estão nos quadros 2 e 3. Numa primeira análise podemos referir que encontrámos alguns resultados interessantes. Vamos apenas realçar as variáveis com valores diferentes em relação às regras de interpretação enunciadas nas obras de Exner (Exner, 2003; Exner e Sendin, 1995).

Nas localizações o valor de Dd é muito elevado, a média é de 5.27 (DP = 4.37, mediana = 4.00, moda = 3.00) e isto é explicado por um número muito elevado de localizações parciais que não constam das listas de Exner. O valor de S também está relativamente elevado, (média = 2.26, DP = 2.22, mediana = 2.00, moda = 2.00) mas não será esta primeira leitura dos resultados que nos poderá permitir concluir que a população portuguesa tem características acentuadas de oposição em relação ao meio.

Quadro 2 - Estatísticas Descritivas para 309 Adultos Não Pacientes

VARIÁVEL	MÉDIA	DP	MIN	MAX	FREQ	MEDIANA	MODA	SK	KU
IDADE	37.25	13.29	18.00	76.00	309	35.00	26.00	0.57	-0.51
ANOS ESC	9.33	4.23	1.00	18.00	309	9.00	4.00	0.22	0.94
R	22.07	7.88	14.00	61.00	309	20.00	16.00	1.74	4.13
W	7.02	4.20	0.00	28.00	303	6.00	5.00	0.97	1.93
D	9.78	5.60	0.00	32.00	307	9.00	8.00	1.10	1.50
Dd	5.27	4.37	0.00	28.00	300	4.00	3.00	1.98	5.61
SPACE	2.26	2.22	0.00	14.00	249	2.00	2.00	1.88	5.04
DQ+	5.15	3.39	0.00	19.00	297	5.00	3.00	0.93	1.23
DQO	15.90	6.58	4.00	46.00	309	15.00	12.00	1.33	3.14
DQV	0.89	1.35	0.00	8.00	137	0.00	0.00	1.94	4.31
DQv/+	0.13	0.38	0.00	2.00	35	0.00	0.00	3.06	9.24
FQX+	0.23	0.84	0.00	11.00	43	0.00	0.00	8.05	91.06

VARIÁVEL	MÉDIA	DP	MIN	MAX	FREQ	MEDIANA	MODA	SK	KU
FQXO	10.51	3.82	2.00	28.00	309	10.00	10.00	0.77	1.35
FQXU	8.37	4.64	1.00	30.00	309	7.00	6.00	1.63	4.18
FQX-	2.72	1.95	0.00	11.00	279	2.00	2.00	1.14	2.15
FQXNONE	0.25	0.58	0.00	3.00	60	0.00	0.00	2.82	8.83
MQ+	0.10	0.41	0.00	3.00	23	0.00	0.00	4.60	23.35
MQO	1.61	1.42	0.00	7.00	231	1.00	1.00	0.87	0.49
MQU	0.79	1.23	0.00	8.00	135	0.00	0.00	2.53	9.40
MQ-	0.29	0.73	0.00	4.00	55	0.00	0.00	2.80	7.85
MQNONE	0.04	0.27	0.00	3.00	10	0.00	0.00	7.32	60.95
SQual-	0.54	0.92	0.00	8.00	116	0.00	0.00	3.29	17.94
M	2.84	2.60	0.00	15.00	258	2.00	1.00	1.38	2.19
FM	3.62	2.87	0.00	19.00	284	3.00	2.00	1.94	6.64
m	1.34	1.51	0.00	7.00	192	1.00	0.00	1.26	1.07
FC	1.25	1.52	0.00	10.00	188	1.00	0.00	1.80	4.48
CF	1.80	1.80	0.00	10.00	224	1.00	0.00	1.24	1.55
C	0.26	0.57	0.00	3.00	61	0.00	0.00	2.34	5.27
CN	0.02	0.19	0.00	3.00	3	0.00	0.00	13.94	210.70
FCCFCCN	3.32	2.76	0.00	15.00	271	3.00	2.00	1.20	1.52
WSUMC	2.81	2.45	0.00	13.00	271	2.00	0.00	1.22	1.46
SUMC'	1.32	1.53	0.00	7.00	180	1.00	0.00	1.24	1.22
SUMT	0.68	0.80	0.00	4.00	154	0.00	0.00	1.06	0.74
SUMV	0.59	1.03	0.00	6.00	109	0.00	0.00	2.49	7.60
SUMY	1.20	1.54	0.00	8.00	169	1.00	0.00	1.58	2.49
SUMSHD	3.78	3.33	0.00	16.00	265	3.00	4.00	1.26	1.64
FR+RF	0.35	0.91	0.00	5.00	55	0.00	0.00	3.02	9.41
FD	1.46	1.48	0.00	9.00	222	1.00	1.00	1.69	4.39
F	9.62	4.58	1.00	28.00	309	9.00	7.00	0.87	1.12
PAIR	7.73	4.82	0.00	32.00	303	7.00	7.00	1.50	4.34
EGO	0.39	0.18	0.00	1.14	307	0.39	0.33	0.51	1.00
LAMBDA	1.21	1.71	0.06	19.00	309	0.75	0.50	6.11	51.99
FM+m	4.96	3.54	0.00	24.00	297	4.00	3.00	1.75	5.28
EA	5.65	4.20	0.00	24.50	296	4.50	4.50	1.38	2.39
ES	8.74	5.42	0.00	32.00	307	8.00	8.00	1.27	2.21
DTOTAL	-0.89	1.46	-7.00	3.00	171	0.00	0.00	-1.07	1.54
ADJD	-0.43	1.21	-7.00	4.00	151	0.00	0.00	-0.82	3.88
a(ACTIVE)	5.06	3.69	0.00	22.00	297	4.00	2.00	1.47	3.14
p(PASSIVE)	2.79	2.51	0.00	19.00	265	2.00	2.00	1.97	7.29
Ma	1.78	1.85	0.00	9.00	224	1.00	0.00	1.44	2.17
Mp	1.09	1.35	0.00	7.00	169	1.00	0.00	1.46	2.29
INTELLCT	2.24	2.58	0.00	15.00	217	2.00	0.00	1.89	4.52
ZF	10.56	4.82	1.00	34.00	309	10.00	10.00	0.84	1.92
ZD	-1.13	4.47	-13.50	16.00	297	-1.00	-0.50	-0.13	0.84
BLENDS	3.53	3.00	0.00	15.00	274	3.00	1.00	1.19	1.39
BLNDS/R	0.16	0.12	0.00	0.67	274	0.14	0.00	0.97	1.04
Col-Shd Bld	0.66	1.05	0.00	7.00	122	0.00	0.00	2.18	6.72
AFR	0.56	0.23	0.19	1.63	309	0.53	0.50	1.02	1.67
POPS	4.93	1.82	0.00	11.00	308	5.00	5.00	0.19	0.04
X+%	0.50	0.13	0.19	0.88	309	0.50	0.50	0.10	-0.19
F+%	0.49	0.19	0.00	1.00	305	0.50	0.50	0.15	0.20
X-%	0.12	0.07	0.00	0.31	279	0.12	0.00	0.29	-0.41
Xu%	0.37	0.12	0.06	0.74	309	0.36	0.33	0.31	0.24
S-%	0.20	0.31	0.00	1.00	116	0.00	0.00	1.51	1.16
Isolate/R	0.22	0.17	0.00	0.94	276	0.19	0.00	0.83	0.65

VARIÁVEL	MÉDIA	DP	MIN	MAX	FREQ	MEDIANA	MODA	SK	KU
H	1.90	1.83	0.00	10.00	248	1.00	1.00	1.63	3.29
(H)	0.79	0.97	0.00	5.00	161	1.00	0.00	1.47	2.44
HD	1.25	1.70	0.00	14.00	189	1.00	0.00	3.12	15.69
(Hd)	0.31	0.69	0.00	5.00	70	0.00	0.00	2.96	11.39
HX	0.24	0.72	0.00	6.00	43	0.00	0.00	4.04	20.16
All H Cont	4.26	3.45	0.00	22.00	290	4.00	3.00	1.98	5.88
A	8.59	3.70	2.00	29.00	309	8.00	6.00	1.06	2.50
(A)	0.21	0.49	0.00	3.00	53	0.00	0.00	2.54	6.60
AD	2.67	2.32	0.00	21.00	275	2.00	1.00	2.60	14.04
(Ad)	0.07	0.29	0.00	2.00	21	0.00	0.00	4.06	17.33
AN	1.38	1.68	0.00	11.00	195	1.00	0.00	1.96	5.38
ART	1.53	1.65	0.00	9.00	204	1.00	0.00	1.44	2.83
AY	0.25	0.61	0.00	4.00	56	0.00	0.00	2.95	9.94
BL	0.25	0.53	0.00	3.00	62	0.00	0.00	2.23	4.71
BT	1.50	1.61	0.00	10.00	212	1.00	0.00	1.74	4.76
CG	1.24	1.47	0.00	9.00	184	1.00	0.00	1.57	3.53
CL	0.26	0.57	0.00	3.00	61	0.00	0.00	2.20	4.19
EX	0.18	0.53	0.00	3.00	41	0.00	0.00	3.32	11.73
FI	0.45	0.82	0.00	6.00	95	0.00	0.00	2.48	8.58
FOOD	0.24	0.50	0.00	3.00	63	0.00	0.00	2.21	5.02
GEOG	0.29	0.78	0.00	7.00	59	0.00	0.00	4.27	24.87
HH	0.65	0.94	0.00	5.00	132	0.00	0.00	1.80	3.77
LS	0.81	1.12	0.00	7.00	144	0.00	0.00	1.79	4.09
NA	0.85	1.18	0.00	6.00	141	0.00	0.00	1.47	1.78
SC	0.67	1.17	0.00	6.00	114	0.00	0.00	2.22	5.05
SX	0.43	0.93	0.00	6.00	79	0.00	0.00	2.88	9.62
XY	0.22	0.66	0.00	6.00	47	0.00	0.00	4.55	27.08
IDIO	0.61	0.92	0.00	6.00	124	0.00	0.00	1.86	4.61
DV	0.31	0.58	0.00	3.00	77	0.00	0.00	1.84	2.79
INCOM	0.31	0.67	0.00	5.00	72	0.00	0.00	2.80	10.61
DR	0.14	0.47	0.00	4.00	33	0.00	0.00	4.25	22.52
FABCOM	0.27	0.53	0.00	3.00	74	0.00	0.00	2.07	4.81
DV2	0.00	0.06	0.00	1.00	1	0.00	0.00	17.58	309.00
INC2	0.00	0.06	0.00	1.00	1	0.00	0.00	17.58	309.00
DR2	0.00	0.00	0.00	0.00	0	0.00	0.00	----	----
FAB2	0.01	0.08	0.00	1.00	2	0.00	0.00	12.37	151.97
ALOG	0.07	0.26	0.00	1.00	22	0.00	0.00	3.35	9.29
CONTAM	0.00	0.00	0.00	0.00	0	0.00	0.00	----	----
SUM6	1.12	1.19	0.00	6.00	200	1.00	1.00	1.37	2.26
LVL2	0.01	0.14	0.00	2.00	3	0.00	0.00	11.89	152.20
WSUM6	2.88	3.37	0.00	20.00	200	2.00	0.00	1.66	3.57
AB	0.23	0.64	0.00	5.00	49	0.00	0.00	3.91	19.11
AG	0.59	0.95	0.00	7.00	122	0.00	0.00	2.57	9.64
CFB	0.00	0.06	0.00	1.00	1	0.00	0.00	17.58	309.00
COP	0.91	1.02	0.00	6.00	179	1.00	0.00	1.43	3.02
CP	0.04	0.22	0.00	2.00	10	0.00	0.00	6.40	44.36
MOR	1.25	1.33	0.00	7.00	202	1.00	0.00	1.39	2.23
PER	1.11	1.30	0.00	7.00	183	1.00	0.00	1.45	2.26
PSV	0.14	0.36	0.00	2.00	42	0.00	0.00	2.31	4.11

As variáveis que compõem o EA são relativamente baixas tanto o M (média = 2.84, DP = 2.60, mediana = 2.00, moda = 1.00), como a Soma de C (WsumC; média = 2.81, DP = 2.45, mediana = 2.00, moda =

0.00). Na Soma C, o CF (média = 1.80) é mais elevado que FC (média = 1.25). Nesta amostra 59% dos sujeitos têm um valor de CF+C mais elevado que FC.

A Soma T é relativamente baixa (média = 0.68), havendo 50% dos sujeitos com Soma T = 0 e 15% com Soma T > 1. A Soma Y (média 1.20) e Soma V (média 0.59) são relativamente elevadas.

O valor de Lambda (média = 1.21, DP = 1.71, mediana = 0.75, moda = 0.50) está bastante elevado em relação ao valor limite de 0.99, observando-se que 111 sujeitos, 36% da amostra, têm um Lambda > 0.99.

O EA (média = 5.65, DP = 4.20) está relativamente baixo e pode ser explicado pelos valores mais elevados de Lambda e valores médios mais baixos em M e Wsum C.

As respostas Popular (P) têm uma média de 4.93 (DP = 1.82), que é um valor relativamente baixo. Verifica-se também que 68 sujeitos (22%) têm $P < 4$ e apenas 7% (22 sujeitos) apresentam um valor de $P > 7$.

As variáveis relacionadas com a Qualidade Formal têm valores bastante baixos. Todavia, e uma vez que a cotação foi realizada segundo as tabelas americanas (Exner, 2001) só teremos uma conclusão mais sólida depois de terminada a elaboração da tabela de Qualidade Formal para a população portuguesa.

Nos conteúdos, o valor de H é baixo (média = 1.90, DP = 1.83, mediana = 1.0, moda = 1.0). Verificamos também que 160 sujeitos (52%) têm valores de H Puro < 2 e 20% da amostra têm mesmo H Puro = 0.

Os valores de COP (média = 0.91, DP = 1.02) e de AG (média = 0.59, DP = 0.95) também são bastante baixos. O valor de MOR (média = 1.25, DP = 1.33) é relativamente elevado.

Outro resultado que nos surpreendeu foi o valor positivo do CDI que se verifica em 48% da amostra ($n = 148$) tendo o CDI=4 em 29% da amostra e o CDI=5 em 19%. Outro resultado interessante que vai necessitar de mais investigação é o DEPI que é positivo em 32% ($n=100$) da amostra (DEPI=5: 19%, DEPI=6: 12%, DEPI=7: 1%).

Se compararmos os resultados por Género as diferenças são pouco significativas. Apenas aparecem diferenças em 6 variáveis. Com $p < 0.01$ temos 2 conteúdos: Ay e Ex. Com $p < 0.05$ temos D, (2), Soma T e Afr (Pires, 2000).

A tentativa de explicar as diferenças entre os resultados da amostra de adultos portugueses e os resultados obtidos por Exner (2001), levou-nos a constatar que o nível de escolaridade é umas das variáveis demográficas mais importantes.

Quadro 3 - Frequências e Percentagens para 33 Variáveis Estruturais - 309 Adultos Não Pacientes

ESTILO EB		QUALIDADE FORMAL	
Introversivo	.62 20%	X+% > .89	.0 0%
Pervasive	.35 11%	X+% < .70	.287 93%
Ambigüal	.87 8%	X+% < .61	.246 80%
Extratensivo	.49 16%	X+% < .50	.146 47%
Pervasive	.31 10%	F+% < .70	.276 89%
Evitativo	.111 36%	Xu% > .20	.284 92%
		X-% > .15	.96 31%
		X-% > .20	.42 14%
		X-% > .30	.2 1%
DIFERENÇAS EA - es: NOTAS-D		PROPORÇÃO FC:CF+C	
Nota D > 0	.23 7%	FC > (CF+C) + 2	.15 5%
Nota D = 0	.138 45%	FC > (CF+C) + 1	.36 12%
Nota D < 0	.148 48%	(CF+C) > FC+1	.104 34%
Nota D < -1	.82 27%	(CF+C) > FC+2	.65 21%
Nota Adj D > 0	.38 12%	S-Constelação Positiva	.20 6%
Nota Adj D = 0	.158 51%	HVI Positivo	.9 3%
Nota Adj D < 0	.113 37%	OBS Positivo	.3 1%
Nota Adj D < -1	.50 16%		
Zd > +3.0 (Overincorp)	.47		
Zd < -3.0 (Underincorp)	.91		
SCZI = 6 0 0%	DEPI = 7 2 1%	CDI = 5 58 19%	
SCZI = 5 0 0%	DEPI = 6 38 12%	CDI = 4 90 29%	
SCZI = 4 1 0%	DEPI = 5 60 19%		
VARIÁVEIS DIVERSAS			
Lambda > .99	.111 36%	(2AB+Art+Ay) > 5	.66 21%
Dd > 3	.180 58%	Populares < 4	.68 22%
DQv + DQv/+ > 2	.40 13%	Populares > 7	.22 7%
S > 2	.100 32%	COP = 0	.130 42%
Soma T = 0	.155 50%	COP > 2	.19 6%
Suma T > 1	.47 15%	AG = 0	.187 61%
3r+(2)/R < .33	.107 35%	AG > 2	.12 4%
3r+(2)/R > .44	.108 35%	MOR > 2	.44 14%
Fr + rF > 0	.55 18%	Nível 2 Sp.Sc. > 0	.3 1%
C Puro > 0	.61 20%	Suma 6 Sp. Sc. > 6	.0 0%
C Puro > 1	.16 5%	H Puro < 2	.160 52%
Afr < .40	.79 26%	H Puro = 0	.61 20%
Afr < .50	.128 41%	% p > a+1	.31 10%
(FM+m) < Soma Shading	.99 32%	Mp > Ma	.69 22%

Nível de escolaridade

No quadro 4, o nível 1 corresponde ao nível de educação básico (0-6 anos ou 0-9 para menores de 30 anos de idade), o nível 2 corresponde ao nível secundário (7-12 anos ou 10-12 anos para menores de 30 anos de idade)

e o nível 3 é o nível superior (+13 anos) que na grande maioria dos nossos sujeitos corresponde à titularidade de um diploma universitário.

O nível básico tem $n=137$ (44% da amostra), o nível secundário $n=111$ (36%) e o nível superior $n=61$ (20%).

Nas 41 variáveis analisadas no quadro 4, existem diferenças significativas em 30 variáveis, tendo com $p<0.05$: FM+m, Fr+rF, Ex, Fi, AG e com $p<0.001$: R, W, S, M, FC, CF, Wsoma C, Soma T, Soma C', Soma V, Soma Y, Lambda, EA, es, FD, a-total activo, p-total passivo, Combinações de determinantes (*Blend*), Combinações cor-esbatido (*Color shading blends*), Popular, H puro, Todos conteúdos H, AB, COP e MOR.

Nestas variáveis com diferença significativa, o nível básico apenas tem o resultado mais elevado na variável Lambda (média = 1.73) em que o nível secundário tem uma média de 0.90 e o nível superior tem a média = 0.58. Também tem os resultados mais elevados em T=0 (61%) e CDI>3 (67%).

Em todas as restantes variáveis o nível secundário tem sempre resultados intermédios pertencendo os resultados mais elevados ao nível de escolaridade superior.

Assim o nível 3 tem R (média = 25.57), W (média = 9.39), S (média = 3.28), M (média = 5.07), FC (média = 2.13), CF (média = 2.87), Wsoma C (média = 4.55), FM+m (média = 5.82), Soma T (média = 1.02), Soma C' (média = 2.08), Soma V (média = 1.16), Soma Y (média = 1.79), EA (média = 9.61), es (média = 11.87), Fr+rF (média = 0.49), FD (média = 1.90), a-total activo (média = 7.20), p-total passivo (média = 3.75), Combinações (*Blends*) (média = 6.13), Combinações cor-esbatido (média = 1.28), Popular (média = 5.66), H puro (média = 2.98), Todos conteúdos H (média = 6.64), Ex (média = 0.26), Fi (média = 0.64), AB (média = 0.67), COP (média = 1.46), AG (média = 1.02), MOR (média = 1.70) e os valores de DEPI>4 (48%), CF+C>FC (41%).

De salientar que os resultados do nível 3 (13 e + anos de educação formal) são os mais interessantes, sendo de destacar o Lambda com a média de 0.58. Todas as variáveis relacionadas com os afectos, ideação e variáveis relacionais aparecem com resultados elevados neste grupo de sujeitos, mesmo que o índice positivo de depressão tenha uma frequência relativamente elevada. Verifica-se assim que o índice CDI positivo aumenta com Lambda elevado e o índice DEPI positivo tem um comportamento inverso, ao aumentar quando o Lambda é mais baixo.

Quadro 4 - Comparação entre Níveis de Escolaridade para 41 variáveis.

Variável	Nível Esc. 1 (n=137)		Nível Esc. 2 (n=111)		Nível Esc. 3 (n=61)	
	Média	DP	Média	DP	Média	DP
Anos Educ	5.52	2.02	10.49	1.27	15.80	1.11
R **	20.32	6.68	22.31	7.55	25.57	9.67
W **	5.86	3.50	7.13	4.05	9.39	4.88
Dd	5.36	4.32	5.00	4.02	5.57	5.09
S **	1.74	1.81	2.32	2.09	3.28	2.87
M **	1.80	1.76	2.92	2.51	5.07	2.95
FC **	0.90	1.33	1.20	1.41	2.13	1.78
CF **	1.22	1.36	1.92	1.80	2.87	2.13
C	0.16	0.46	0.29	0.56	0.41	0.74
WSOMA C **	1.91	1.72	2.95	2.35	4.55	3.00
FM+m *	4.40	3.32	5.17	3.07	5.82	4.54
SOMA T **	0.47	0.65	0.76	0.81	1.02	0.96
SOMA C' **	0.82	1.17	1.50	1.65	2.08	1.81
SOMA V **	0.34	0.78	0.58	1.01	1.16	1.32
SOMA Y **	0.78	1.28	1.40	1.73	1.79	1.48
Lambda **	1.73	2.33	0.90	0.86	0.58	0.39
EA **	3.71	2.53	5.87	4.02	9.61	4.65
ES **	6.81	4.42	9.40	4.98	11.87	6.48
Fr+rF *	0.24	0.77	0.41	0.95	0.49	1.07
EGO	0.39	0.18	0.39	0.19	0.40	0.16
FD **	1.12	1.32	1.64	1.67	1.90	1.26
ACTIVO **	4.01	2.75	5.20	3.60	7.20	4.67
PASSIVO **	2.23	2.34	2.95	2.08	3.75	3.20
AFR	0.54	0.21	0.59	0.25	0.57	0.22
BLENDS **	2.15	1.91	3.81	2.84	6.13	3.42
ColShblends **	0.32	0.64	0.75	0.98	1.28	1.52
P **	4.39	1.72	5.21	1.67	5.66	1.94
X+%	0.49	0.13	0.52	0.13	0.49	0.13
X-%	0.13	0.08	0.12	0.07	0.11	0.07
XU%	0.37	0.12	0.35	0.12	0.39	0.12
H **	1.26	1.36	2.10	1.85	2.98	2.12
TODOS CONT H**	3.11	2.53	4.36	3.00	6.64	4.62
A	8.46	3.72	8.83	3.63	8.44	3.83
Bl	0.23	0.54	0.20	0.42	0.38	0.66
Ex *	0.11	0.43	0.23	0.59	0.26	0.60
Fi *	0.30	0.78	0.53	0.74	0.64	1.00
Food	0.20	0.46	0.22	0.45	0.34	0.65
AB **	0.07	0.28	0.19	0.61	0.67	0.99
COP **	0.62	0.84	0.96	1.03	1.46	1.13
AG *	0.46	0.82	0.50	0.84	1.02	1.26
MOR **	0.93	1.21	1.40	1.27	1.70	1.53
DEPI>4	22%		37%		48%	
DEPI>5	9%		14%		22%	
T=0	61%		45%		34%	
CDI>3	67%		41%		19%	

(*): $p>0.05$

(**): $p>0.001$

Conclusão

Numa primeira análise, a nível do total da amostra, os resultados deste estudo normativo mostram que a nossa população apresenta resultados muito diferentes quando comparados com os dados normativos de Exner. Esta constatação poderia permitir concluir que as diferenças seriam devidas à influência cultural.

Quando analisamos os resultados tendo em conta as variáveis demográficas, comparando os resultados dos diferentes subgrupos, constatou-se que os níveis escolares tinham uma importância fundamental, para explicar os resultados do nosso estudo. Com o aumento do nível de escolaridade, variáveis como Lambda, EA, WSoma C e Soma T, passam a ter resultados semelhantes aos do estudo normativo de Exner, nos E.U.A. A análise da amostra de adultos americanos, permite constatar que a maioria dos sujeitos tem mais de 12 anos de escolaridade, o que corresponde ao nosso grupo educacional de nível 3. Temos então variáveis que são sensíveis à influência educacional e que são explicadas por aspectos intraculturais.

Por outro lado, verifica-se a existência de outras variáveis, como as relacionadas com a qualidade formal, com resultados muito estáveis em todos os grupos de escolaridade da nossa amostra e que apresentam valores muito diferentes, quando comparados com os resultados de Exner. Os resultados deste tipo de variáveis mostrariam a importância da vertente intercultural.

Nos vários estudos normativos, realizados em vários países de diferentes continentes, quando se encontram diferenças significativas, estas são explicadas através das diferenças culturais, nomeadamente, as diferenças linguísticas, o nível de desenvolvimento tecnológico, a organização social, as diferenças sócio-culturais. A hipótese de as diferenças sócio-culturais poderem ter alguma influência no funcionamento psicológico dos indivíduos, é a justificação principal, implícita ou explícita, de todos os estudos normativos do Teste de Rorschach.

Este estudo mostra que a diferença cultural continua a ser importante, mas realça o valor da variável nível de escolaridade, que passa a explicar um grande número de diferenças de resultados entre amostras de países ou culturas diferentes. Não podemos deixar de mencionar que o sujeito além de realizar uma “leitura perceptiva” dos cartões Rorschach, tem que exprimir verbalmente aquilo que percepcionou. A escolaridade média e longa ajudam a desenvolver a capacidade de expressão verbal do sujeito e facilitam uma postura mais confortável na situação de aplicação do teste. Podemos também levantar a hipótese de que uma escolaridade longa poderá ter, ao nível do fun-

cionamento psicológico do indivíduo, um maior impacto do que a maioria dos psicólogos habitualmente pensa.

Estes resultados indicam também a necessidade da análise de cada protocolo ser cada vez mais precisa, tendo em conta os dados normativos específicos do grupo a que o sujeito pertence, tal como Exner (2003) referiu relativamente aos estilos de resposta em vez de se utilizar os resultados da amostra total, porque estes são demasiado heterogéneos.

Referências

- Campo, V., e Vilar, N. (1999, Julho). *Demythologization of the non-patient concept as a normative ideal*. Comunicação oral apresentada no XVI International Congress of Rorschach and Projective Methods, Amsterdam, Holanda.
- Ephraim, D., Riquelme, J., e Occupati, R. (1996). Características psicológicas de habitantes de Caracas según el Sistema Compreensivo del Rorschach. Datos normativos y comparación transcultural. In D. Ephraim (Ed.), *El método Rorschach en la actualidad* (pp. 45-66). Caracas, Venezuela: Monte Ávila Editores Latinoamericana.
- Exner, J. E. e Sendin, C. (1995). *Manual de interpretación del Rorschach Para el Sistema Compreensivo*. Madrid: Psimática.
- Exner, J. E. (1993). *The Rorschach: A Comprehensive System. Volume 1: Basic foundations* (3.^a ed.). New York: Wiley.
- Exner, J. E. (2001). *A Rorschach Workbook for the Comprehensive System*. (5.^a ed.). Asheville: Rorschach Workshops.
- Exner, J. E. (2002). A new nonpatient sample for the Rorschach Comprehensive System: A progress report. *Journal of Personality Assessment*, 78(3), 391-404.
- Exner, J. E. (2003). *The Rorschach: A Comprehensive System. Volume 1: Basic foundations and principles of interpretation* (4.^a ed.). New York: Wiley.
- INE (1993). *Censos 91: Resultados Definitivos*. Lisboa: Instituto Nacional de Estatística.
- INE (2002). *Censos 2001: Resultados Definitivos*. Lisboa: Instituto Nacional de Estatística.
- Mattlar, C. E., Carlsson, A., Forsander, C., Norrlund, I., Oist, A. S., Maki, J., e Alanen, E. (1995). Rorschach content and structural summary features characteristic of adult, in cross-cultural comparison. In V. Moita e A. A. Pires, (Eds.), *XIV Congresso Internacional de Rorschach e Métodos Projectivos: Proceedings Book*. (pp. 213-220). Lisboa: Sociedade Portuguesa de Rorschach.
- Mormont, C. (1998, Agosto). *Données normatives en Belgique pour le Rorschach en Système Intégré*. Comunicação oral apresentada no 5.º Congresso da ERA (European Rorschach Association), Madrid.
- Pires, A. A. (1992, Outubro). *O estudo normativo do teste de Rorschach: Aspectos metodológicos*. Comunicação oral apresentada no III Simpósio Nacional de Investigação em Psicologia, organizado pela Associação Portuguesa de Psicologia na Fundação Gulbenkian, Lisboa, Portugal.
- Pires, A. A. (1994, Julho). *The Rorschach normative study in Portugal: First results and comparison with the Exner normative study in the U.S.A.* Comunicação oral apresentada na 5th Annual Convention of the International Council of Psychologists, Lisboa, Portugal.
- Pires, A. A. (1995). O estudo normativo do teste de Rorschach na população portuguesa: Resultados parciais. In V. Moita e A. A. Pires (Eds.), *XIV Congresso Internacional*

de Rorschach e Métodos Projectivos: Proceedings Book. (pp. 233-240). Lisboa: Sociedade Portuguesa de Rorschach.

- Pires, A. A. (1996, Julho). *The Rorschach normative study in Portugal: Cross-cultural comparison with the U.S.A. normative data*. Comunicação oral apresentada no XV International Congress of Rorschach and Projective Methods, Boston, E. U. A..
- Pires, A. A. (1996, Novembro). *O estudo normativo do Rorschach na população portuguesa: Comparação intercultural com os dados normativos dos E.U.A.* Comunicação oral apresentada na IV Conferência Internacional “Avaliação Psicológica: Formas e Contextos”, Braga, Portugal.
- Pires, A. A. (1998, Agosto). *The Rorschach normative study in Portugal*. Comunicação oral apresentada no V Congreso da E.R.A. - (European Rorschach Association), Madrid, Espanha.
- Pires, A. A. (1999, Julho). *The Rorschach normative study in Portugal*. Comunicação oral apresentada no XVI International Congress of Rorschach and Projective Methods, Amsterdam, Holanda.
- Pires, A. A. (2000). National norms for the Rorschach Comprehensive System in Portugal. In R. Dana (Ed.), *Handbook of cross-cultural and multicultural personality assessment* (pp. 367-392) Mahwah, NJ: Lawrence Erlbaum Associates.
- Pires, A. A. (2001, Março). *The Rorschach normative study in Portugal: Methodological issues and results*. Symposium: An International Symposium on Rorschach non-patient data: Worldwide Findings. Comunicação oral apresentada no 2001 SPA Mid-winter Meeting, Philadelphia, E.U.A..
- Pires, A. A. (2002, Setembro). *The Rorschach normative study (adults) in Portugal: The coping styles results*. Comunicação oral apresentada no XVII International Congress of Rorschach and Projective Methods, Roma, Itália.
- Pires, A. A. (2004, Agosto). *Methodological issues of the Rorschach normative study (adults) in Portugal*. Special Symposium on Methodological Issues in Normative Data Collection. Comunicação oral apresentada no VII ERA- European Rorschach Association, Congress, Stockholm, Suécia.
- Pires, A. A. (2004, Agosto). *Why so many high Lambda results*. Special Symposium on Methodological Issues in Normative Data Collection. Comunicação oral apresentada no VII ERA- European Rorschach Association, Congress, Stockholm, Suécia.
- Pires, A. A. (2005, Julho). *The Popular Responses, D and Dd Locations in the adult normative portuguese population*. Comunicação oral apresentada no XVIII International Congress of Rorschach and Projective Methods, Barcelona, Espanha.
- Sendín, M. C. (1995). Non-patient transcultural comparison. In V. Moita, e A. A. Pires (Eds.), *XIV Congresso Internacional de Rorschach e Métodos Projectivos: Proceedings Book* (pp. 207-212). Lisboa: Sociedade Portuguesa de Rorschach.
- Shaffer, T. W., Erdberg, P., e Haroian, J. (1999). Current nonpatient data for the Rorschach, WAIS-R and MMPI-2. *Journal of Personality Assessment*, 73(2), 305-316.

- Silva, D. R. (1991) Análise do Rorschach de três grupos de crianças portuguesas no termo do 1º Ano de escolaridade. *Revista Portuguesa de Psicologia*, 27, 61-73.
- Silva, D. R., e Prazeres, N. (1994) Sobre dados de alguns estudos normativos do Sistema Integrativo do Rorschach (SIR) de Exner fora dos E.U. *Psicologica*, 12, 37-53.
- Silva, D. R., Campos, R. and Dias, A. M. (2001) A normative study with portuguese children aged six to ten years old. *SIS Journal of Projective Psychology e Mental Health*, 8 (2), 107-115.
- Silva, D. R., e Dias, M. (2004). Cognição e nível intelectual no Rorschach de crianças dos 6 aos 10 anos. *Revista Iberoamericana de Diagnóstico y Evaluación Psicológica*, 18 (2), 79-97.
- Silva, D. R., Novo, R., e Prazeres, N. (1990). Los datos normativos del Rorschach presentados por Exner son válidos para la Población europea *in extenso*? Ensayo con una población portuguesa. *Revista de la Sociedad Española del Rorschach y Métodos Projectivos*, 3, 33-40.
- Silva, D. R., Novo, R., e Prazeres, N. (1991). Serão os dados normativos do Rorschach apresentados por Exner válidos *in extenso* para a população europeia? Ensaio com uma população portuguesa (versão integral). *Revista Portuguesa de Psicologia*, 27, 13-27.
- Silva, D. R., Novo, R., e Prazeres, N. (1992). Diferenças entre crianças portuguesas e americanas em algumas variáveis do Rorschach: Uma questão transcultural. *Revista Portuguesa de Psicologia*, 28, 149-162
- Silva, D. R., e Campos, R. (2000). Some Rorschach variables from a normative study with 10-year old portuguese children. In A. Andronikof-Sanglade (Ed.), *Rorschachiana Yearbook of the International Rorschach Society*, 24, 110-126. Seattle, Hogrefe e Huber Publishers.
- Silva, D. R., e Prazeres, N. (1995). Sobre dados de alguns estudos normativos do Sistema Integrativo do Rorschach (SIR) de Exner fora dos E.U.. *Revista Ibero-Americana de Diagnóstico y Evaluación Psicológica*, 2, 161-187.
- Silva, D. R., e Campos, R. (1999). Questões referentes a algumas variáveis do Rorschach suscitadas por um estudo normativo com crianças portuguesas de dez anos. *Psicologica*, 22, 71-81.
- Silva, D. R., Novo, R., e Prazeres, N. (1996). The evolution of some Rorschach variables in portuguese children. *European Journal of Psychological Assessment*, 12 (1), 53-58.
- Sultan, S., Andronikof, A., Fouques, D., Lemmel, G., Mormont, C., Reveillere, C., Saias, T. (2004). Vers des normes francophones pour le Rorschach en systeme integre: Premiers resultats sur un echantillon de 146 adultes. *Psychologie Francaise*, 49 (1), 7-24.
- Vinet, E., Saiz, J. L., e San Martín, C. (1995). Necesidad de normas nacionales en el Sistema Comprensivo del Psicodiagnóstico de Rorschach: El caso de Chile. *Revista Iberoamericana de Diagnóstico y Evaluación Psicológica*, 2, 189-201.

THE RORSCHACH NORMATIVE STUDY IN THE ADULT PORTUGUESE POPULATION: METHODOLOGICAL ISSUES AND SOME RESULTS

António Abel Pires

Faculdade de Psicologia e Ciências da Educação, Universidade do Porto

Abstract: This research work presents some results of the Rorschach Normative Study in the Portuguese adult population, with the Comprehensive System.

The results permits differentiate between the cross-cultural and intra-cultural orientation of many Rorschach variables. We can identify Rorschach variables that are sensitive to the cross-cultural dimension, and their results are stable in the three educational levels but different of the Exner data in the U.S.A. (Exner, 1993). There are other Rorschach variables that are sensitive to the differences inside the same culture, which changes according to the educational level. It is noted that the increase of the educational level accounts to the disappearance of some differences between the results of our entire sample and the U.S.A. Normative Study.

KEY-WORDS: *Rorschach test, Normative study, Adults, Portugal.*

NOMEAÇÃO DE FACES FAMOSAS: CAPACIDADE DE EVOCACÃO DE NOMES PRÓPRIOS NUMA AMOSTRA POPULACIONAL PORTUGUESA

Isabel Pavão Martins, Clara Loureiro,
Susana Rodrigues, Beatriz Dias
*Laboratório de Estudos de Linguagem, Unidade Neurológica de Investigação Clínica,
IMM, Faculdade de Medicina de Lisboa*

Resumo

A dificuldade em evocar nomes próprios é uma queixa frequente na população idosa e em doentes com patologia cerebral. Descrevemos um teste de nomeação de faces famosas, desenhado para a população portuguesa, constituído por 71 fotografias de personagens públicas de diferentes nacionalidades, épocas e categorias profissionais. Foi aplicado a 250 sujeitos, entre os 18 e os 84 anos (média 41.2 anos), 64.4% dos quais do sexo feminino. Verificámos um marcado efeito de literacia nesta prova, mas não de género ou idade. A consistência interna do teste foi de .95. O teste revelou-se específico para rastreio de lesão cerebral, quando avaliado numa amostra clínica. Este instrumento pode ser utilizado na avaliação neuropsicológica de indivíduos com queixas cognitivas ou com lesão cerebral.

PALAVRAS-CHAVE: *Nomes próprios, nomeação, evocação de nomes, prosopnomia.*

Introdução e Fundamentação Teórica

Os nomes próprios têm um papel fundamental na comunicação humana e a incapacidade em evocar os nomes dos outros pode causar embaraço social ou ofensa. Talvez por isso, a dificuldade em aceder aos nomes próprios é uma queixa frequente na população adulta saudável, sobretudo nos indivíduos idosos (Lovelace e Twohig, 1990; Evrard, 2002).

Vários estudos indicam que os nomes próprios são particularmente difíceis de aprender (Cohen, 1990) e de evocar (Evrard, 2002), o que tem sido atribuído à fragilidade das associações aos seus referenciais únicos. De facto, os nomes próprios são considerados expressões referenciais puras (Semenza e Zettin, 1989), uma vez que, ao contrário dos nomes comuns, não lhes estão

Morada (address): Prof. Dr. Isabel Pavão Martins. Laboratório de Estudos de Linguagem, Centro de Estudos Egas Moniz, Faculdade de Medicina de Lisboa, 1649-028 Lisboa. Telf./Fax: 217934480, Email: labling@fm.ul.pt

inerentes quaisquer atributos ou propriedades intrínsecas, e a sua associação aos indivíduos é única e inteiramente arbitrária.

Existe igualmente evidência de que os processos cognitivos utilizados na evocação dos nomes próprios se podem segregar, no sistema nervoso, de uma via mais geral, utilizada na nomeação de outros estímulos (nomeadamente objectos manipuláveis e animais), podendo inclusivamente corresponder a uma localização neuroanatômica específica (Damásio e colaboradores, 1996). Esta hipótese baseia-se, não só na descrição de casos com defeitos específicos de evocação de nomes próprios, mas igualmente de casos de anomia, que poupam os nomes próprios, constituindo duplas dissociações entre os nomes próprios e os nomes comuns (McKenna e Warrington, 1978; Farrajota e colaboradores, 1999; Lyons, Hanley e Kay, 2002). Esta teoria tem sido corroborada por estudos de lesão e estudos de activação cerebral, que demonstram a existência de áreas selectivamente activadas na evocação (Damásio e colaboradores, 1996) ou preferencialmente activadas na aprendizagem (Herholz e colaboradores, 2001) dos nomes próprios.

Existe alguma controvérsia sobre quais os factores responsáveis por esta segregação dos nomes próprios, nas redes neuronais utilizadas no acesso e evocação lexical. O facto de serem nomes particularmente difíceis, quando comparados com os nomes comuns, não parece ser suficiente para explicar esta dissociação, uma vez que existem descrições de defeitos gerais de nomeação que poupam os nomes próprios (McKenna e colaboradores, 1978; Farrajota e colaboradores, 1999). Enquanto alguns autores defendem que a diferenciação entre nomes próprios e nomes comuns depende das características semânticas dos estímulos (Crutch e Warrington, 2004), outros evocam a maior complexidade perceptiva dos estímulos, assim como os seus referenciais únicos, para explicar essa segregação.

Os Testes de Nomeação de Faces Famosas (TFF) constituem o instrumento mais comum para avaliar a evocação de nomes próprios. Têm sido desenvolvidos, essencialmente, para caracterizar defeitos de nomeação e de memória semântica, em doentes amnésicos ou com demência e para estudar casos de prosopagnosia. Mais recentemente, têm sido também utilizados para avaliar doentes com lesões focais do sistema nervoso e indivíduos com suspeita de deterioração cognitiva.

Na literatura são diversos os TFF descritos, alguns aferidos em populações saudáveis e outros de carácter experimental (Sanders e Warrington, 1971; Albert, Butters e Levin, 1979; Stevens, 1979; McKenna e Warrington, 1980; Faglioni, Cremonini e De Renzi, 1991; Pluchon e colaboradores, 2002; Rizzo, Venneri e Papagno, 2002). Não temos conhecimento da existência de TFF pa-

ra a população portuguesa. Embora já tivéssemos desenvolvido versões experimentais deste tipo de prova, no estudo de casos clínicos (Farrajota e colaboradores, 1999), não dispúnhamos de dados normativos para este tipo de tarefa na nossa população. Dada a frequência desta queixa na clínica, a sua sensibilidade ao envelhecimento normal e o seu contributo diagnóstico em fases iniciais da demência de Alzheimer (Brazzelli e colaboradores, 1994; Delazer e colaboradores, 2003), sentimos a necessidade de desenvolver um instrumento que nos permitisse quantificar e caracterizar esse defeito.

O objectivo deste trabalho foi, por conseguinte, o desenvolvimento de um TFF para a população Portuguesa. Este teste foi desenhado de forma a ser aplicável a indivíduos adultos, com mais de 18 anos de idade, de ambos os sexos e de qualquer nível de escolaridade ou proveniência social, desde que o Português fosse a sua primeira língua e que os indivíduos estivessem familiarizados com a cultura contemporânea portuguesa. Pretendíamos ainda que pudesse ser utilizado em indivíduos saudáveis ou em doentes com lesão cerebral, se os mesmos tivessem boa capacidade visual e de compreensão das instruções do teste.

O TFF consiste na apresentação de fotografias de personagens públicas. O sujeito deve produzir o respectivo nome, ou eventualmente descrever os seus atributos, da forma mais correcta possível. O teste, que aqui descrevemos, foi apresentado em reuniões científicas ao longo do seu processo de construção e de obtenção de dados normativos (Martins e colaboradores, 2002), embora nunca na sua versão definitiva.

Quanto às dimensões avaliadas nesta prova, sabe-se que esta tarefa envolve, habitualmente, vários processos cognitivos, activados sequencialmente e em cascata: a) percepção visual e reconhecimento da face como familiar (nível pré-semântico); b) identificação do estímulo, por activação de um conjunto de memórias de longo termo que lhe estão associadas (cargo, profissão, actividades, idade, nacionalidade ou se está vivo ou não) (nível semântico); e c) activação lexical, com procura, escolha e evocação da forma fonológica do nome e sua produção (nível lexical ou pós-semântico). Dificuldades em qualquer destas etapas podem perturbar a capacidade de nomeação. Pode-se, eventualmente, identificar o nível do defeito através de provas adicionais (emparelhamento de faces, reconhecimento de faces pelo nome, nomeação por outras modalidades sensoriais ou evocação de atributos) e, igualmente, pela análise das respostas e dos erros (nomeadamente a presença do fenómeno "tip-of-the-tongue" ou "debaixo da língua") ou dos mecanismos facilitadores da evocação.

Método

Amostra

A colheita dos dados normativos na população saudável foi desenhada como um estudo observacional, transversal e prospectivo. Todos os indivíduos da amostra participaram na base do voluntariado, através do consentimento escrito ou oral (nos iletrados). Os dados foram colhidos em dois períodos distintos: estudo piloto e estudo definitivo. Os dados colhidos durante o estudo piloto não foram utilizados no estudo definitivo.

A selecção dos indivíduos foi baseada num plano de amostragem estratificado por idade e escolaridade, de forma a assegurar a representação dos vários níveis etários e educacionais. Assim, foi constituída uma grelha com 6 grupos de idade (por décadas, desde a segunda até à sétima) e 3 grupos de escolaridade, que se tentaram preencher de forma equilibrada. Contudo, os dados relativos a alguns subgrupos foram difíceis de obter, nomeadamente dos indivíduos mais jovens de baixa escolaridade e dos idosos de elevado nível educacional. Esta amostra, de conveniência, foi ainda seleccionada através de um questionário para rastreio de lesão encefálica. Os critérios de inclusão no estudo foram os seguintes: a) idade igual ou superior a 18 anos; b) ausência de história passada ou presente de lesão cerebral (nomeadamente epilepsia e patologia vascular ou traumática do SNC) ou de doença psiquiátrica crónica; c) sem evidência de deterioração mental (confirmada por uma pontuação acima do ponto de corte no Mini Mental State Evaluation (MMSE) (Folstein, Folstein e McHugh, 1975); e d) ausência de toxicod dependência, alcoolismo ou ingestão de psicofármacos.

Material

Concepção e construção do teste

O desenho do TFF, assim como a análise dos seus resultados, basearam-se em testes construídos para outras línguas e culturas, nomeadamente para a população Inglesa (Sanders e colaboradores, 1971; Stevens, 1979), Norte Americana (Albert e colaboradores, 1979), Italiana (Faglioni e colaboradores, 1991; Rizzo e colaboradores, 2002) e Francesa (Pluchon e colaboradores, 2002).

As faces-estímulo, reproduções fotográficas a preto e branco, em formato 10 x 15cm, de faces de personagens públicas, foram retiradas de revistas, jor-

nais, livros ou da Internet, colocadas em álbuns e apresentadas individualmente, sempre pela mesma ordem.

Os itens foram seleccionados de forma a incluir personagens nacionais e internacionais, de diferentes sectores da vida pública (política/governo, entretenimento/vida social, cultura e desporto) e de diferentes épocas (no seu pico de popularidade). De facto, uma vez que existe uma gradação temporal no conhecimento das personagens públicas, pretendíamos, posteriormente, utilizar essa gradação como medida temporal da memória semântica. Desta forma, organizaram-se as personagens por épocas correspondentes ao seu pico de popularidade, de modo a que o seu reconhecimento pudesse constituir um índice de memória referente a uma determinada época da vida. Assim, foram seleccionadas, por consenso entre os autores, 200 fotografias para o estudo piloto, tendo em consideração as seguintes características: a) fotografias nítidas, facilmente reconhecíveis, em que a face se apresentasse de frente e sem sombras e, de preferência, da época correspondente ao seu pico de popularidade; b) ausência de adereços identificadores (ex. coroas, fardas).

Seleção dos itens

O estudo piloto teve como objectivo a selecção dos estímulos para a prova definitiva, e consistiu na avaliação de 50 indivíduos (20 homens e 30 mulheres) de idades compreendidas entre os 20 e os 75 anos, distribuídos por 5 grupos etários.

Os resultados do estudo piloto foram analisados item a item. Das 200 imagens apresentadas, foram seleccionadas as nomeadas por mais de 70% dos sujeitos (nome completo ou parte mais relevante do nome e nenhum elemento incorrecto), uma vez que se pretendia assegurar a exequibilidade do teste na população de baixa escolaridade. Verificou-se, no entanto, que, seguindo este critério, se obtinha um conjunto enviesado de 55 fotografias, com um predomínio das personagens contemporâneas mais mediáticas (jogadores de futebol e personagens televisivas), deixando, por conseguinte, de representar as diferentes épocas e sectores da sociedade desejados. Assim, e de forma a assegurar a presença de estímulos representativos de todos os sectores profissionais e períodos de tempo, seleccionámos outros itens, entre os mais nomeados, das categorias em falta, obtendo um teste de 74 fotografias, que foram utilizadas na colheita dos dados normativos.

Critérios utilizados na cotação das respostas

A forma de cotar as respostas em provas de denominação de faces públicas, raramente é descrita noutros estudos. No entanto, um mesmo estímulo po-

de ter várias respostas correctas, mais ou menos completas, conforme é evocada parte do nome ou o nome completo. Foi estabelecido, como regra geral, que uma resposta que incluísse a parte nuclear do nome, ou seja, a parte evocada por mais de 90% da população de controlo, seria considerada a resposta correcta e cotada com 1 ponto. Às evocações parciais do nome, respostas dadas por menos de 10% do grupo de controlo, foram atribuídas uma pontuação de zero.

Por exemplo, para o estímulo “Francisco Pinto Balsemão”, mais de 90% da população incluiu os nomes “Pinto Balsemão” na sua resposta, pelo que essa foi considerada a parte nuclear do nome. Respostas parciais como “Francisco” ou “Balsemão”, embora correctas, foram cotadas com zero, por serem incompletas e pouco frequentes na população saudável. Naturalmente todas as respostas que incluíram um ou mais nomes incorrectos (“António Pinto Balsemão”) foram pontuadas como zero. Ainda de acordo com este critério, em alguns itens houve necessidade de aceitar duas respostas correctas, uma vez que se verificou uma distribuição binomial das respostas, nunca se tendo atingido os 90%. Por exemplo, para o mesmo item, 57.7% dos participantes responderam “Charlie Chaplin” e 32% “Charlot”.

Procedimentos

Após a descrição dos objectivos e procedimentos do estudo e a obtenção do respectivo consentimento, foram recolhidos, para todos os indivíduos, os seus dados biográficos (idade, sexo, lateralidade, local de residência, nível educacional) e preenchido um questionário de saúde, de modo a assegurar que cumpriam os critérios de inclusão do estudo. Foram colhidos elementos relativos à zona geográfica de residência (classificada em Urbana ou Rural e em três áreas geográficas - Norte, Centro e Sul, incluindo a zona de Lisboa e Vale do Tejo).

Todos os sujeitos de idade igual ou superior a 60 anos foram também avaliados pela versão Portuguesa do MMSE (Guerreiro e colaboradores, 1994), para excluir deterioração cognitiva. Procedeu-se, de seguida, à aplicação do teste de nomeação de faces.

As fotografias foram organizadas, num único álbum, por ordem crescente de dificuldade, tendo sido a apresentação sequencial (uma a uma), sem tempo limite. Foram usadas as 74 faces seleccionadas no estudo piloto. Foi dada ao sujeito a seguinte instrução “Vou-lhe mostrar uma série de fotografias. Quero que me diga o nome destas pessoas, o mais completo possível. Pode demorar o tempo que quiser”. As respostas foram registadas em

folha própria, tendo sido transcrita a resposta exacta do sujeito. Nos primeiros 103 sujeitos foi, igualmente, efectuada a recolha dos atributos dos fotografados ("Diga-me agora tudo o que sabe sobre cada uma destas pessoas"). Estes dados destinavam-se a esclarecer a capacidade de reconhecimento das faces e o nível semântico do processamento do estímulo, independentemente da sua nomeação.

Os participantes foram avaliados em diferentes locais, nomeadamente no Hospital de Santa Maria (funcionários do Hospital, familiares e acompanhantes de utentes do Serviço de Neurologia), em Centros de Saúde (acompanhantes de utentes dos Centros de Saúde de Castelo Branco, Azeitão e Évora), em Centros de Dia para idosos (Clínica de São João de Deus) e na Academia da Terceira Idade em Oeiras. Os directores dos Centros de Saúde e Centros de Dia foram formalmente contactados por escrito e informados sobre as características e procedimentos do estudo, tendo sido solicitada a respectiva autorização. O protocolo do estudo foi aprovado pela Comissão de Ética do Hospital de Santa Maria.

As observações foram realizadas por elementos do Laboratório de Estudos de Linguagem e por alunos a efectuarem Estágios de Investigação Clínica da Faculdade de Medicina de Lisboa.

A confidencialidade dos dados foi salvaguardada, não tendo sido incluídos quaisquer elementos de identificação na base de dados. A análise dos resultados foi efectuada através do programa SPSS.

Resultados

Na amostra definitiva, foram observados 250 indivíduos, a maioria do sexo feminino (64.4%), com uma média de idade de 41.2 anos (± 15.4), variando entre os 18 e os 84 anos. A distribuição por faixa etária foi a seguinte: 18-30 anos (28.8%), 31-40 anos (20.4%), 41-50 anos (18.8%), 51-60 (18%), 61-70 anos (8.8%) e acima dos 71 anos (5.2%). A maioria dos indivíduos (65.2%) tinha mais de 10 anos de escolaridade, 20.4% tinham entre 5 e 9 anos de escolaridade e os restantes menos de 5 anos. A maioria dos indivíduos eram dextros (94.8%) e residentes em regiões urbanas (85%) do Sul do país, incluindo Lisboa e Vale do Tejo (83%).

Nos 250 sujeitos da amostra, verificámos não haver diferenças na distribuição entre a idade e o género ($\chi^2=4.02$, $p=n.s.$), nem entre a escolaridade e o género ($\chi^2=1.03$, $p=n.s.$). Observou-se uma relação inversa, significativa, entre a idade e a escolaridade ($r=-.46$, $p<.01$).

Dados qualitativos e quantitativos dos itens

Para analisar o índice de discriminação dos itens, foram efectuadas correlações bisseriais entre os itens e a nota total. Todos os itens, excepto três, apresentaram correlações r s superiores a .20 com a nota total (estatisticamente significativas para $p < .01$). As três excepções (Mário Soares, Herman José e Marco Paulo) constituíam itens muito fáceis, com uma percentagem de respostas correctas de 99.2%, 96.8% e 95.6%, respectivamente, pelo que foram eliminados, reduzindo o número de itens do teste para 71.

O índice de dificuldade dos itens variou entre .14 e .96. Verificou-se que a maioria dos itens (69%) se situava na faixa metrologicamente desejável, entre .20 e .80, tendo sido o valor médio do índice de dificuldade .70, tal como seria de esperar, uma vez que a selecção efectuada no estudo piloto favoreceu as faces com maior percentagem de respostas correctas.

Quanto à análise do funcionamento diferencial dos itens, não foi encontrada diferença significativa de género em 63 dos 74 estímulos. Em 6 itens (4 do desporto), constatou-se um resultado significativamente superior nos indivíduos do sexo masculino, enquanto que noutros 5 (personagens do espectáculo), se verificou o contrário. Uma vez que o número estava equilibrado entre os dois sexos, decidiu-se manter esses itens.

Os itens apresentaram sensibilidade ao nível educacional e, alguns deles, à idade dos examinados, tal como era esperado.

Resultados no âmbito da precisão

O coeficiente de consistência interna relativo à totalidade da amostra foi de .95 (alfa de Cronbach) para os 71 itens seleccionados.

A estabilidade temporal dos resultados foi examinada numa amostra de 30 indivíduos (21 mulheres e 9 homens), com uma média de idades de 39.0 anos (± 14.9), testados em dois momentos distintos do tempo, com um intervalo que variou entre os 30 e os 171 dias (em média 91.1 dias). Comparando a nota obtida nas duas observações, verificou-se uma estabilidade temporal adequada ($r = 0.92$, $p < .01$), tendo sido a nota da segunda avaliação significativamente superior à da primeira, o que demonstra um efeito de aprendizagem. O acordo entre cotadores foi investigado em 48 participantes (27 do sexo feminino) com uma média de idades de 37.8 anos e 14.2 anos de escolaridade. Nestes casos, as respostas de cada indivíduo saudável foram cotadas por dois observadores independentes. O coeficiente de fiabilidade entre cotadores foi elevado ($r = 1.00$, $p < .01$), o que evidencia que o teste é

bastante fácil de cotar, desde que o examinador disponha da grelha das respostas que são consideradas correctas.

Resultados relativos à validade

O teste foi construído com base noutros testes idênticos (validade facial). O número de controlos que participaram no estudo normativo foi superior ao de outros testes publicados.

No que respeita à validade de construto foram utilizadas diferentes medidas. Não dispúnhamos de nenhum teste idêntico, previamente realizado, relativamente ao qual se pudesse determinar a validade convergente. A comparação com testes de nomeação de objectos era pouco útil, uma vez que, em indivíduos saudáveis, se tende a obter um efeito de tecto (Nunes, Guerreiro e Martins, 2000).

O marcado efeito de escolaridade ($F(2,247)=38.4$, $p<.01$) sobre o desempenho nesta prova fornece-nos uma medida de validade por factor externo. O conhecimento de personagens públicas é uma medida de memória semântica, dependente do nível educacional do indivíduo.

Tal como já foi referido, o TFF avalia três domínios cognitivos gerais. Estudámos a correlação desta prova com quatro tarefas destinadas à avaliação desses domínios, para efectuar a validade concorrente do TFF. Assim, aplicámos simultaneamente numa amostra de sujeitos, o TFF e um teste de percepção visual ($N=30$) (Line Orientation Test de Benton) (Benton e colaboradores, 1983) (nível pré semântico), a prova de reconhecimento das mesmas faces, pela descrição dos seus atributos ($N=103$) (nível semântico) e duas provas de fluência verbal, semântica e fonémica para a letra P ($N=30$) (nível lexical). Verificou-se uma correlação significativa com qualquer um deles ($r_s=.65$, $p<.01$; $r_s=.56$, $p<.01$; $r_s=.61$, $p<.01$; $r=.92$, $p<.01$, respectivamente).

Efectuámos ainda uma análise factorial exploratória, com extracção de componentes principais usando uma Rotação Varimax no conjunto das 74 faces. Foram extraídos 4 factores com valor próprio (*eigenvalues*) superior a 1, cuja interpretação é tentativa. Esses factores poderão representar diferentes gradações da memória semântica no tempo e diferentes interesses por categorias profissionais. Todos os factores se relacionaram de forma estreita com a escolaridade ($p<.01$). Esta análise irá ser detalhada num estudo posterior.

Normas, critérios e parâmetros

Considerando a amostra estudada ($N=250$) e a versão de 71 itens, a pontuação média obtida foi de 49.5 ± 15.5 (variando entre 1 e 71) com um

erro padrão de medida (Epm) de .92. Verificou-se uma correlação significativa entre a escolaridade e a pontuação total ($r=.53$, $p<.01$). Uma análise de variância (ANOVA) mostrou diferenças significativas entre os três grupos de escolaridade considerados ($F(2,25)=38.4$, $p<.01$). Através da análise de comparações múltiplas de Tukey HSD, verificou-se que os três grupos de escolaridade apresentavam diferenças significativas entre si, em qualquer par de comparações (Figura 1).

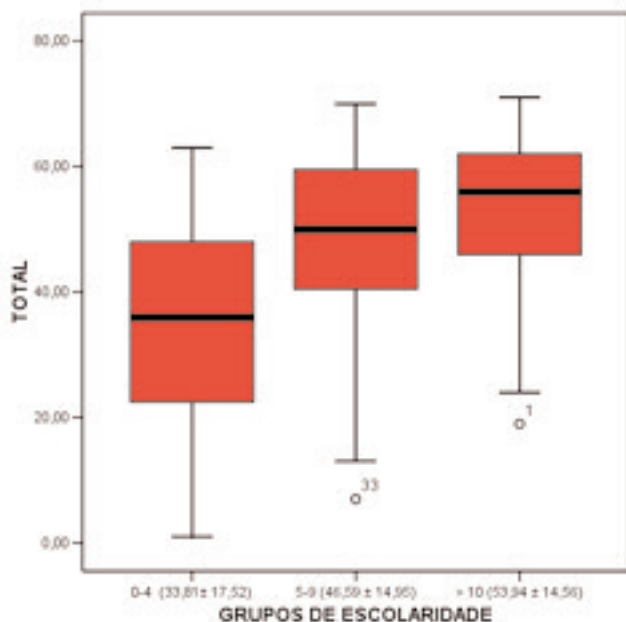


Figura 1 - Pontuação média obtida no teste de acordo com os grupos de escolaridade. As diferenças entre os grupos são significativas em qualquer par de comparações ($p<.01$).

Relativamente à idade, não foi encontrada uma correlação significativa com a pontuação na prova ($r=-.09$, $p=.15$, n.s.), mesmo controlando o factor escolaridade (coeficiente de correlação parcial=.08, $p=.29$, n.s.). Contudo, quando a idade foi categorizada por décadas, verificou-se uma diminuição significativa do desempenho nos participantes mais idosos (> 70 anos) ($F(5)=2.6$, $p=.03$). Uma vez que o nosso grupo de sujeitos idosos é reduzido, e enviesado no sentido de uma baixa escolaridade, pensamos que será importante estudar este aspecto numa amostra populacional maior.

Não foram encontradas diferenças significativas na pontuação total para os outros parâmetros avaliados (sexo, lateralidade, residência urbana ou rural e região Norte, Centro ou Sul do país).

Estudo da sensibilidade e especificidade do teste numa amostra clínica

O TFF foi aplicado a uma amostra de conveniência, constituída por doentes com lesão cerebral conhecida, observados no Laboratório de Estudos de Linguagem, por motivos clínicos (N=75). Trata-se de uma amostra heterogênea de doentes, que inclui doentes com afasia ligeira (N=9), nos estadios iniciais da doença de Alzheimer (N=4) ou demência vascular (N=9), indivíduos com defeito cognitivo ligeiro (N=17) e um grupo de doentes com lesão cerebral, sem evidência de afasia, nem demência, com lesões cerebrovasculares (N=8), epilepsia temporal por esclerose mesial (N=21), doença de Parkinson (N=2), tumor intracraniano (N=1) e traumatismos cranioencefálicos (N=4).

O Quadro 1 indica, as médias das notas padrão ("t scores", ou seja, notas já corrigidas para a escolaridade, sendo a média 50 e o desvio padrão 10), obtidas pelos diferentes grupos clínicos. Verificámos que os afásicos são o grupo com mais baixo desempenho nesta prova (t médio = 25.1), seguindo-se os doentes com demência, dos quais os casos de demência de Alzheimer (t médio = 30.7) obtiveram valores mais baixos, em média, que os doentes com o diagnóstico de demência vascular (t médio = 33.9).

Quadro1 - Pontuações em nota padrão (sendo a média 50 e o desvio padrão 10) obtidas pela amostra clínica, por diagnóstico.

Diagnóstico	N	Média	Desvio padrão
Demência	13	32.97	7.30732
MCI	17	34.81	16.55632
Afasia	9	25.06	12.54872
Outros	36	35.86	11.69926
Total	75	33.83	12.70429

Comparámos ainda este grupo clínico com um grupo controle, emparelhado por idade e escolaridade, para avaliar a sensibilidade e especificidade do teste no rastreio de disfunção cerebral. Categorizámos a pontuação obtida na prova em dois grupos: resultados normais (sempre que a pontuação obtida fosse igual ou superior a -1.5 desvios padrões, relativamente à média para o respectivo grupo de escolaridade, isto é, "t score" >35) e resultados patológicos (abaixo desse valor). A partir desta categorização, calculámos a sensibilidade e especificidade do teste através do programa Confidence Interval Analysis (2000). Verificámos, nesta amostra que o teste tinha uma especificidade de .92 (95%CI .84-.96), sensibilidade de .45 (95%CI .35-.57), um valor preditivo positivo de .85 (95%CI .72-.93), valor preditivo negativo de .63 (95%CI .54-.71) e uma razão de verosimilhança de 5.83 (95%CI 2.6 -1.3) na identificação de disfunção cerebral.

Discussão e avaliação crítica

O teste de nomeação de faces é um teste de aplicação muito simples e rápida, entre 10 a 20 minutos, de acordo com as capacidades do examinado. Tem como principal vantagem a sua adequação à população portuguesa, com normas para diferentes níveis de escolaridade. Permite-nos, assim, avaliar a ausência ou o grau de dificuldade na evocação de nomes próprios. Os itens estão organizados por grau crescente de dificuldade, permitindo interromper a prova se houver uma grande dificuldade nos itens iniciais, evitando a frustração do examinado.

As principais limitações desta prova relacionam-se com: a) a sua previsível desactualização, com necessidade de introdução de estímulos mais recentes; e b) a necessidade de actualizar os dados normativos.

Reconhecemos ainda que o grupo de controlo é uma amostra de conveniência obtida na base do voluntariado, podendo não ser representativo da população adulta portuguesa. Reconhecemos que estão particularmente pouco representados nesta amostra os indivíduos jovens de baixa escolaridade (difíceis de encontrar entre os controlos, devido à escolaridade obrigatória) e os idosos de alta escolaridade.

No que respeita à sua utilidade clínica, pensamos que é um teste útil para caracterizar os defeitos de nomeação em doentes com vários tipos de patologia e que, associado ao teste de reconhecimento das mesmas faces, pode ser também utilizado para estudar a memória semântica e as suas gradações temporais. Os resultados obtidos no grupo clínico, estão de acordo com a nossa expectativa e com estudos publicados. De facto, a dificuldade de nomeação é dos sinais cardinais da afasia e o desempenho de doentes afásicos, mesmo com afasias ligeiras, tenderá a ser inferior ao normal numa tarefa de nomeação considerada difícil, como esta. Em segundo lugar, estão os doentes com demência de Alzheimer em estadio inicial. Seguem-se doentes com outros tipos de patologia do sistema nervoso central. É preciso notar que alguns dos grupos incluídos podem ter pontuações dentro da média, uma vez que a patologia subjacente não afecta as tarefas cognitivas envolvidas nesta prova. De facto, pensamos que a fraca sensibilidade encontrada nesta amostra clínica reflecte esta heterogeneidade de patologias e de desempenhos na prova. A sua aplicação em grupos clínicos mais homogêneos, nomeadamente na afasia e na demência de Alzheimer, faz parte dos projectos que pretendemos desenvolver no futuro. Na realidade, estudos realizados noutras populações mostraram que este teste pode ser particularmente útil no diagnóstico precoce da demência de Alzheimer.

Encontram-se em curso vários projectos que poderão expandir a aplicação clínica deste teste:

a) Análise dos atributos por épocas. Pretendemos com essa análise, desenvolver uma prova de memória semântica e de reconhecimento de faces públicas, que poderá complementar o presente teste, permitindo o diagnóstico diferencial entre prosopagnomia, prosopagnosia e defeitos semânticos;

b) Estudo do valor ecológico do teste, em populações idosas com queixas de dificuldades de evocação de nomes próprios;

c) Estudo do desempenho de doentes com diversas patologias clínicas, nomeadamente epilepsia temporal, defeito cognitivo ligeiro e demência de Alzheimer;

d) Elaboração de duas versões breves alternativas da prova, com idêntico grau de dificuldade.

Referências

- Albert, M. S., Butters, N., e Levin, J. (1979). Temporal gradients in retrograde amnesia of patients with alcoholic Korsakoff's disease. *Archives of Neurology*, 36, 211-216.
- Altman D.G. Machin e Bryant T. Gardner. (2000). *Statistics with confidence 2nd Ed.*. Confidence Interval Analysis v 20.0. BMJ books.
- Benton, A. L., Hamsher, K. S., Varney, N. S., e Spreen, O. (1983). Judgement of Line Orientation, form H. *Contributions to Neuropsychological Assessment*. Oxford University Press, NewYork, NY.
- Brazzelli, M., Capitani, E., Della Sala, S., Spinnler, H., e Zuffi, M. (1994). A neuropsychological instrument adding to the description of patients with suspected cortical dementia : the Milan overall dementia assessment. *Journal of Neurology, Neurosurgery and Psychiatry*, 57(12), 1510-7.
- Cohen, G. (1990). Why is it difficult to put names to faces? *British Journal of Psychology*, 81, 287-297.
- Crutch, S. J., e Warrington, E. K. (2004). The semantic organisation of proper nouns: the case of people and brand names. *Neuropsychologia*, 42(5), 584-596.
- Damásio, H., Grabowski, T. J., Tranel, D., Hichwas, R. D., e Damásio, A. R. (1996). A neural basis for lexical retrieval. *Nature*, 380, 499-505.
- Delazer, M., Semenza, C., Reiner, M., Hofer, R., e Benke, T. (2003). Anomia for people names in DAT – evidence for semantic impairments. *Neuropsychologia*, 41(12), 1593 -1598.
- Evrard, M. (2002). Ageing and Lexical Access to Common and Proper Names in Picture Naming. *Brain and Language*, 81, 174-179.

- Faglioni, P., Cremonini, A. M., e De Renzi, E. (1991). Taratura su soggetti normali di test di facce sconosciute e familiari. *Archivio di Psicologia, Neurologia e Psichiatria*, 52, 339-350.
- Farrajota, L., Mariano, G., Baeta, E., Canhão, P., e Pavão-Martins, I. (1999). Proper name anomia : a double dissociation. *Journal of International Neuropsychological Society*, 5(4), 275.
- Folstein, M., Folstein, S., e McHugh, P. (1975). "Mini mental state": a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3), 189-198.
- Guerreiro, M., Silva, A. P., Botelho, M. A., Leitão, O., Castro Caldas, A., e Garcia C. (1994). Adaptação à população portuguesa da tradução do "Mini Mental State Examination (MMSE)". *Revista Portuguesa de Neurologia*, suppl 1.
- Herholz, K., Ehlen, P., Kessler, J., Strotmann, T., Kalbe, E., e Markowitsch, H-J. (2001). Learning face-name associations and the effect of age and performance: a PET activation study. *Neuropsychologia*, 39, 643-650.
- Lovelace, E. A., e Twohig, P. T. (1990). Healthy older adult's perceptions of their memory functioning and use of mnemonics. *Bulletin of the British Psychological Society*, 28, 115-118.
- Lyons, F., Hanley, J. R., e Kay, J. (2002). Anomia for common names and geographical names with preserved retrieval of names of people: a semantic memory disorder. *Cortex*, 38(1), 23-35.
- Martins, I. P., Loureiro, C., Gameiro, H., Miguel, C., Martins, V., e Alves, A. (2002). Naming Public Faces: a Test for the Portuguese Population (preliminary results). *Journal of International Neuropsychological Society*, 8(4), 486.
- McKenna, P., e Warrington, E. K. (1978). Category-specific naming preservation: A single case study. *Journal of Neurology, Neurosurgery and Psychiatry*, 41, 571-574.
- McKenna, P., e Warrington, E. K. (1980). Testing for nominal dysphasia. *Journal of Neurology, Neurosurgery and Psychiatry*, 43, 781-788.
- Nunes, M., Guerreiro, M., e Martins, I. P. (2000). Naming and Literacy: The Influence of Age. *Journal of the International Neuropsychological Society*, 6, 400.
- Pluchon, C., Simonnet, E., Toullat, G., e Gil, R. (2002). Les effets du vieillissement normal sur la dénomination des visages et la reconnaissance de personnages célèbres: batterie 75. *Revue Neurologique (Paris)*, 158(6-7), 703-708.
- Rizzo, S., Venneri, A., e Papagno, C. (2002). Famous face recognition and naming test: a normative study. *Neurological Sciences*, 23, 153-159.
- Sanders, H. I., e Warrington, E. K. (1971). Memory for remote events in amnesic patients. *Brain*, 94, 661-668.
- Semenza, C., e Zetlin, M. (1989). Evidence from aphasia for the role of proper names as pure referring expressions. *Nature*, 342, 678-679.
- Stevens, M. (1979). Famous personality test. A test for measuring remote memory. *Bulletin of the British Psychological Society*, 32, 211.

Agradecimentos:

Os autores agradecem a colaboração: a) de todos os voluntários que se disponibilizaram para participar no estudo e dos Directores dos centros de Saúde e Centros de Dia, que autorizaram a sua realização; b) dos alunos da Faculdade de Medicina de Lisboa, que participaram na colheita de parte dos dados, no âmbito dos Estágios de Investigação Clínica: Ana Margarida Alves, Cláudia Miguel, Vera Martins, Hugo Gameiro e Sandra Barreto; c) da Dra. Luisa Albuquerque, Professor Alexandre Mendonça, Terapeutas da Fala Gabriela Leal, José Fonseca e Luisa Farrajota, assim como às Dras. Filipa Ribeiro e Sandra Ginó, pela disponibilidade demonstrada na validação clínica da prova; d) da Dra. Raquel Gil Gouveia, pelo apoio prestado no tratamento estatístico dos dados; e) da Dra Sara Ramos pela ajuda fornecida na recolha da bibliografia.

NAMING FAMOUS FACES PROPER NAME RETRIEVAL IN A SAMPLE OF THE PORTUGUESE POPULATION

Isabel Pavão Martins, Clara Loureiro,
Susana Rodrigues, Beatriz Dias

Laboratório de Estudos de Linguagem, UNIC, IMM, Faculdade de Medicina de Lisboa

Abstract: Difficulty in proper name retrieval is a frequent cognitive complaint among elderly subjects and brain damaged patients. The purpose of this study was to develop a test of naming famous faces for the Portuguese population. Healthy participants (N=250), whose age ranged from 18 to 84 years, were asked to name 71 photographs of public faces, from different nationalities, occupation and epochs. There was a marked literacy effect on the final score but no gender or age differences. The internal consistency of the test was .95, and it proved to be quite specific to brain pathology, when applied to a clinical sample. This test may be useful for the clinical evaluation of subjects with brain dysfunction or with cognitive complaints.

KEY-WORDS: *Proper names, naming, name retrieval, prosopagnosia.*

AUTO-AVALIAÇÃO DA MEMÓRIA DO QUOTIDIANO: ALGUMAS QUESTÕES RELATIVAS AO USO DE QUESTIONÁRIOS DE METAMEMÓRIA E ESTUDO DAS PROPRIEDADES PSICOMÉTRICAS DO QUESTIONÁRIO DE FUNCIONAMENTO DA MEMÓRIA (QFM)*

Maria Salomé Pinho
Mário R. Simões

Faculdade de Psicologia e de Ciências da Educação, Universidade de Coimbra

Resumo

A utilização de questionários para avaliar o funcionamento da memória do quotidiano levanta várias problemas quanto à qualidade das medidas, entre as quais se destacaram a falta de consistência e o “paradoxo da introspecção da memória”. Algumas das variáveis que influenciam as respostas neste tipo de questionários foram indicadas considerando a perspectiva da cognição social. Dada a importância do conhecimento relativo às propriedades psicométricas das pontuações das auto-avaliações da memória, analisaram-se os resultados da aplicação do Questionário de Funcionamento da Memória (QFM), numa amostra de 427 estudantes universitários, no que diz respeito à fiabilidade (consistência interna e estabilidade temporal teste-reteste), validade de constructo (análise em componentes principais), validade interna (intercorrelações entre as escalas do QFM) e correlações com notas escolares. De um modo geral, a versão portuguesa do QFM apresentou boas propriedades psicométricas. No entanto, verificou-se que a solução factorial encontrada difere no agrupamento dos itens correspondentes a dois dos quatro factores originais.

PALAVRAS-CHAVE: *Memória, metamemória, questionários, fiabilidade, validade.*

Introdução

O Questionário de Funcionamento da Memória (QFM; *Memory Functioning Questionnaire*, MFQ) de Gilewski, Zelinski e Schaie (1990) insere-se na abordagem que se baseia no modo como as próprias pessoas avaliam as suas capacidades e/ou desempenhos mnésicos no dia-a-dia, i. e., numa me-

* Investigação financiada pela Fundação Calouste Gulbenkian (Procº 63537). Uma versão deste estudo foi apresentada na VII European Conference on Psychological Assessment, Málaga, 1-4 Abril de 2004.

Morada (address): Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra. Rua do Colégio Novo, Apartado 6153. 3001-802 Coimbra. Portugal. Email: salome@fpce.uc.pt

metodologia típica de auto-avaliação da memória em tarefas do quotidiano¹. Mais especificamente, este questionário reporta-se, quase em exclusivo, à dimensão da metamemória, relacionada com o conhecimento de como funciona a memória e com o uso de estratégias comportamentais na execução de tarefas que envolvam competências mnésicas, que foi denominada, por McDougall (1998), “conhecimento factual”².

A informação recolhida a partir da consciência (ou do estar ciente) do funcionamento mnésico levanta alguns problemas, o que origina, por vezes, uma completa rejeição deste tipo de metodologia. Baddeley (1999) refere, a propósito deste debate, uma citação retirada de um diário de um estudante de direito, que ficou amnésico na sequência de uma hemorragia cerebral:

«Late last night I remembered that I'd forgotten to list things that I had forgotten. But then how I know what I'd forgotten?» (p. 313)

A aparência paradoxal desta situação não deverá esconder uma questão situada num nível mais fundamental e que se prende com a dificuldade em dispor, em qualquer etapa, de procedimentos independentes do observador ou cujas condições de influência sejam conhecidas e controladas de forma rigorosa. A insuficiência dos resultados dos estudos sobre a complexidade da recuperação de informação na memória tem também aqui um papel determinante.

No âmbito da sua investigação sobre a fiabilidade de avaliações subjectivas, baseadas em questionários para a identificação de défices de memória, Baddeley (1999) sublinha que é necessário conhecer, justamente, de onde provém a falta de consistência encontrada neste tipo de avaliação. Relativamente a este aspecto, há a considerar os efeitos resultantes do estado de funcionamento cognitivo e das atitudes da própria pessoa que responde ao questionário, bem como a adequação do instrumento aplicado.

A avaliação feita através de questionários e/ou listas de queixas mnésicas, também designada “avaliação subjectiva”, não deverá ser a única a ter em conta na averiguação das funções mnésicas. Mesmo a abordagem baseada em resultados obtidos em tarefas laboratoriais tem problemas associados que, não raras vezes, são remetidos, apressadamente, para a esfera da validade dita ecológica. Contudo, para autores como, por exemplo, Chow (1987) trata-se de uma questão mal colocada, uma vez que a validade ecológica reporta-se à teoria e não às experiências ou tarefas experimentais aplicadas.

¹ Para a enumeração e descrição das metodologias de auto-avaliação da memória, veja-se Simões e Ferreira (1996; p. 184).

² As outras dimensões da metamemória destacadas por McDougall (1998) aparecem também referidas em Pinho e Simões (no prelo).

Voltando ao estudo de Baddeley (1999), na sua conclusão é afirmado que os questionários e as listas de queixas de memória podem ser úteis na obtenção de informação qualitativa sobre os aspectos que mais perturbam uma pessoa. Porém, estes não fornecem indicações fiáveis da gravidade do défice.

O uso de questionários e/ou listas de queixas de memória, quer para avaliar o desempenho mnésico quer para determinar o efeito de crenças sobre a memória na execução de tarefas mnésicas, requer o conhecimento de variáveis que influenciam as respostas dadas pelos sujeitos e cuja necessidade foi já mencionada. Neste contexto, um dos trabalhos de Herrmann (1982) ocupa uma posição singular. As crenças acerca do funcionamento da memória, que estão envolvidas em ambas as utilizações deste tipo de instrumentos, apresentam um carácter relativamente estável, as respostas obtidas variam com o tipo de dificuldade mnésica, o grau em que a pessoa é cognitivamente afectada quando se encontra sob *stress*, a confiança nas suas competências mnésicas e a sua idade. Todavia, as próprias crenças ou, de um modo mais genérico, a metamemória não é muito acurada (Herrmann, 1982). Se saber o que uma pessoa pensa, acerca do funcionamento das suas capacidades mnésicas, é relevante na tentativa de promover a adequação entre potencialidades cognitivas percebidas e exigências das tarefas a desempenhar (Pinto, 1997), é igualmente necessário ter presente que estratégias decorrentes das crenças podem revelar-se inadequadas devido ao carácter inapropriado das primeiras. Não menos importante, é ainda considerar alguma heterogeneidade quando se fala na estabilidade daquilo que as pessoas pensam sobre a sua experiência passada. Com efeito, há crenças mais estáveis do que outras (p. ex., a fiabilidade da recordação de situações de esquecimento relativas a conversas é menor do que se se tratar de esquecimento de lugares) e, por outro lado, algumas pessoas não são estáveis na sua avaliação da memória (Herrmann, 1982). Esta ausência de estabilidade poderá, em parte, dever-se ao declínio da própria memória, colocando-se aqui, novamente, o «paradoxo da introspecção da memória», segundo a designação seguida por Herrmann – como pedir a uma pessoa com competência mnésica deficitária para recordar como é que *verdadeiramente* funciona a sua memória?

Um outro ponto de vista, sobre as variáveis que afectam as respostas obtidas em questionários e listas de queixas, que envolvem auto-avaliações da própria memória, provém da abordagem em termos de cognição social. A especificidade desta perspectiva reside no tipo de conteúdos considerado, quando se afirma que a formação das respostas é baseada na informação armazenada que se torna acessível. Estes conteúdos são relativos a componentes do auto-conceito ou, de um modo mais geral, a construções mentais acerca de

si próprio e dos outros. A este propósito ir-se-ão destacar, brevemente, as explicações decorrentes da adopção de dois conceitos: *esquematicidade* (Markus, 1977) e *teorias implícitas* (Ross, 1989).

Um indivíduo é classificado como esquemático se relativamente a um conceito do tipo traço, por exemplo “boa capacidade mnésica”, considera esse conceito como um bom auto-descritor e como sendo muito importante (Cavanaugh *et al.*, 1998). Estudos sobre a relação entre *esquematicidade* e crenças sobre a memória revelaram que adultos mais velhos apresentam níveis mais elevados de respostas afectivas em relação às falhas de memória do que adultos mais jovens. O que foi interpretado como um aumento da *esquematicidade*, no que diz respeito à memória e à perda de competências mnésicas com o avanço da idade. Da *esquematicidade* decorre um processamento com características similares às encontradas nos peritos num dado domínio como, por exemplo, a captação automática de informação relevante, organização da informação na memória em constructos facilmente acessíveis, maior elaboração na integração de informação consistente, influência acrescida das atitudes sobre o comportamento, em virtude da ocorrência automática de avaliações e da menor propensão para ser afectado por efeitos do contexto. De modo mais concreto, é importante assinalar que as respostas de auto-avaliação da memória são ainda mais influenciáveis pelo formato e tipo de formulação dos itens do questionário, quando se apela para aspectos da memória relativamente aos quais um indivíduo não é considerado esquemático (*ibid.*).

Inerente ao conceito de *teorias implícitas*, encontra-se a ideia segundo a qual a recordação de atributos pessoais, subjacente às respostas de auto-avaliação, supõe o recurso a uma concepção implícita acerca da evolução (estabilidade ou mudança) no que respeita à posição actual nos atributos em causa (*ibid.*). Contudo, a influência das *teorias implícitas* não se apresenta com um carácter obrigatório. Estas nem sempre se encontram acessíveis, surgem como credíveis ou são consideradas relevantes. Por exemplo, a recuperação de informação na memória relativa a eventos específicos ou a pormenores particulares destes depende menos do efeito de *teorias implícitas*. Mas, se a capacidade de processamento se encontrar restringida, devido, por exemplo, à menor focalização da atenção, então o papel das teorias implícitas torna-se mais importante, podendo ocorrer uma das seguintes situações: a informação específica ou particular de experiências passadas recuperáveis é substituída por tais teorias ou a codificação e a inferência com base em informação actualmente recuperada são por estas enviesadas. Assim, ao não se conhecerem as *teorias implícitas* dos sujeitos, acerca da memória e do envelhecimento, as estimações que estes fazem das eventuais mudanças das suas competências

mnésicas podem dificultar a correcta interpretação das respostas obtidas nos questionários (Cavanaugh et al., 1998). Tais teorias, se incluírem a ideia de declínio com o aumento da idade, podem levar a uma auto-avaliação em termos de deterioração (apesar de existir informação proveniente de outras fontes que possa contrariar tal apreciação), mas também a uma sobrestimação das suas capacidades mnésicas no passado (*ibid.*). É ainda relevante saber se as *teorias implícitas* são revistas ao longo da vida do sujeito e quais as condições que desencadeiam a sua influência.

Brevemente, importa sublinhar que nos processos construtivos, que ocorrem durante a recuperação de informação armazenada sobre a própria memória e o *self*, intervêm estratégias gerais de recuperação como, por exemplo, a disponibilidade, a acessibilidade e a representatividade³.

As questões psicométricas relativas às pontuações das auto-avaliações da memória ou das crenças sobre o funcionamento da memória são ainda consideradas prioritárias (Cavanaugh e Baskind, 1996; Cavanaugh et al., 1998). É importante distinguir os aspectos psicométricos mais sólidos daqueles que o não são, para que, relativamente aos primeiros, se explorem outras vertentes do processo de validação e, no que diz respeito aos últimos, seja identificado em que sentido deve prosseguir o aperfeiçoamento do instrumento em análise (Cavanaugh e Baskind, 1996). Neste contexto, ir-se-ão referir os estudos psicométricos por nós realizados com o QFM, sendo este considerado um dos instrumentos de metamemória com melhores características metrológicas (Cavanaugh et al., 1998) e que tem motivado estudos diversos (Hertzog et al., 1989; Hertzog et al., 1990; Lane e Zelinsky, 2003; McDonald-Miszczak e Hertzog, 1995; Zelinski et al., 1990).

Método

Participantes

O QFM foi aplicado numa amostra de 427 participantes (382 pertencentes ao género feminino e 45 ao género masculino), com uma média de idade

³ No contexto da investigação sobre tomada de decisão, a disponibilidade refere-se à facilidade de recuperação, i. e., à influência da informação que se encontra facilmente disponível na formação de uma resposta por parte do sujeito e não deve ser confundida com a disponibilidade mnésica, que remete para a presença ou armazenamento de informação na memória.

A acessibilidade diz respeito à situação em que informação facilmente acessível afecta a resposta do sujeito, ao passo que na investigação sobre memória traduz o grau em que a informação pode ser recuperada.

Quanto à representatividade, o sentido a considerar é o de informação, que se julga representativa do tipo de eventos em causa, influenciar a resposta dada (Ashcraft, 1989; Cavanaugh et al., 1998).

de 19.98 anos (mediana igual a 19 e desvio-padrão de 2.68). Esta amostra apresenta um nível de escolaridade elevado, pois todos os participantes eram estudantes universitários (1º, 2º e 3º anos das licenciaturas de psicologia e de ciências da educação da FPCE-UC).

Instrumento

O QFM resultou de um conjunto de estudos que tiveram por objectivo melhorar o Questionário de Metamemória (QM; Zelinski *et al.*, 1980). Dos 92 itens do QM foram excluídos 28 por não saturarem significativamente nos quatro factores considerados: «frequência geral do esquecimento», «gravidade do esquecimento», «funcionamento retrospectivo» e «uso de mnemónicas» (Gilewski *et al.*, 1990). Assim, o QFM é composto por 64 itens, a cada um dos quais está associada uma escala de 1 a 7 pontos, relativa a juízos de capacidade-qualidade (p. ex., “como está a sua memória comparada com o que era há cinco anos atrás?”) ou a juízos de frequência (p. ex., “enquanto está a ler um romance, quão frequentemente tem problemas em se recordar do que foi lido três ou quatro capítulos antes daquele que se encontra a ler?”). Uma pontuação mais elevada significa, de um modo geral, uma auto-avaliação de melhor funcionamento mnésico no quotidiano.

Procedimento

A versão portuguesa do QFM incluiu os procedimentos habituais de tradução e de retroversão por dois tradutores independentes. Do questionário original foram retirados dois itens («como está a sua memória comparada com o que era há 20 anos atrás?» e «como está a sua memória comparada com o que era quando tinha 18 anos?») porque muito poucos participantes (apenas 3.3%) tinham mais de 24 anos de idade.

O QFM foi administrado em pequenos grupos e, algumas vezes, individualmente, pelos autores deste trabalho.

Resultados

Estudo da fiabilidade

A fiabilidade de um instrumento refere-se à sua consistência e à estabilidade das pontuações ao longo do tempo. A consistência interna foi estimada com base no cálculo do coeficiente alfa de Cronbach, que foi determinado primeiramente para a globalidade do questionário (alfa = .87). As análises

que se seguem tomam já em consideração cada uma das escalas decorrentes da estrutura factorial encontrada (ver adiante a secção “Estudo da validade”).

No quadro 1 são apresentados os valores do coeficiente alfa, os coeficientes de correlação de cada item com o total da escala “Frequência geral do esquecimento e avaliação do funcionamento mnésico” (excluindo o próprio item) e também as médias e desvios-padrão.

Quadro 1 - *Análise da consistência interna da escala “Frequência geral do esquecimento e avaliação do funcionamento mnésico” do QFM*

Itens	Média	Desvio-padrão	Correl.Item/total (excluído o próprio item)	Alfa de Cronbach (excluído o próprio item)
1	4.89	.94	.48	.84
2A	4.58	1.13	.32	.84
2D	4.26	1.19	.39	.84
2E	5.14	1.12	.34	.84
2F	4.30	1.37	.31	.84
2G	4.44	1.40	.45	.84
2H	5.88	1.13	.41	.84
2I	5.04	1.04	.35	.84
2J	4.86	1.15	.34	.84
2K	4.70	1.54	.42	.84
2L	4.93	1.10	.32	.84
2M	5.28	1.22	.34	.84
2O	5.39	1.17	.41	.84
2P	4.92	1.06	.48	.84
2Q	5.05	1.11	.51	.84
2R	4.65	1.20	.50	.84
3A	4.71	1.10	.43	.84
5A	5.07	1.16	.46	.84
5B	4.73	1.07	.51	.84
5C	4.24	1.13	.49	.84
5D	3.58	1.31	.37	.84
7A	4.08	.91	.30	.84
7B	4.09	1.38	.36	.84
7C	4.24	1.69	.36	.84

A inspecção dos valores do quadro 1 revela que a escala “Frequência geral do esquecimento e avaliação do funcionamento mnésico” do QFM apresenta bons indicadores de consistência. O valor do coeficiente alfa para a amostra foi de .85. Todos os itens possuem valores de correlação significativos com o total do questionário e superiores ou iguais a .30.

Apresentam-se a seguir, no quadro 2, os mesmos indicadores para a escala “Frequência do esquecimento durante a leitura”, composta por 9 itens.

Quadro 2 - *Análise da consistência interna da escala "Frequência do esquecimento durante a leitura" do QFM*

Itens	Média	Desvio-padrão	Correl.Item/total (excluído o próprio item)	Alfa de Cronbach (excluído o próprio item)
3B	5.15	1.12	.52	.90
3C	5.87	.93	.69	.88
3D	6.19	.94	.70	.88
3E	6.40	.87	.61	.89
4A	5.13	1.01	.49	.90
4B	5.46	1.00	.69	.88
4C	6.03	.85	.79	.87
4D	6.15	.88	.80	.87
4E	6.36	.81	.70	.88

Com base no quadro 2 pode verificar-se que as correlações de cada item com o total (sem incluir o próprio item) variam entre .49 e .80. O coeficiente alfa de Cronbach para o conjunto dos itens que fazem parte desta escala ("Frequência do esquecimento durante a leitura") foi de .90.

Os valores do coeficiente alfa, das correlações de cada um dos 18 itens com o total da escala "Gravidade do esquecimento" (excluindo o próprio item) e também as médias e desvios-padrão estão inscritos no quadro 3.

Quadro 3 - *Análise da consistência interna da escala "Gravidade do esquecimento" do QFM*

Itens	Média	Desvio-padrão	Correl.Item/total (excluído o próprio item)	Alfa de Cronbach (excluído o próprio item)
6A	4.20	1.42	.40	.89
6B	3.99	1.47	.49	.89
6C	2.88	1.51	.56	.89
6D	3.69	1.38	.50	.89
6E	4.52	1.36	.43	.89
6F	4.00	1.31	.47	.89
6G	4.03	1.63	.43	.89
6H	3.21	1.63	.55	.89
6I	3.29	1.42	.63	.89
6J	4.51	1.33	.47	.89
6K	3.50	1.52	.49	.89
6L	3.86	1.54	.55	.89
6M	3.60	1.73	.57	.89
6N	2.87	1.53	.49	.89
6O	2.83	1.54	.69	.89
6P	3.18	1.49	.68	.89
6Q	2.81	1.45	.67	.89
6R	3.92	1.44	.51	.89

Da análise do quadro 3 pode concluir-se que as correlações de cada item com o total (excluindo o próprio item) apresentam valores pertencentes ao in-

tervalo .40 – .69. O valor de alfa encontrado para a escala “Gravidade do esquecimento” foi de .89.

A escala “Uso de mnemónicas” do QFM é constituída por 8 itens e os respectivos valores de alfa de Cronbach, coeficientes de correlação de cada item com o total da escala (excluindo o próprio item) e também as médias e desvios-padrão podem ser consultados no quadro 4.

Quadro 4 - *Análise da consistência interna da escala “Uso de mnemónicas” do QFM*

Itens	Média	Desvio-padrão	Correl.Item/total (excluído o próprio item)	Alfa de Cronbach (excluído o próprio item)
8A	3.67	1.83	.41	.77
8B	3.25	1.45	.66	.73
8C	3.68	1.73	.66	.72
8D	3.66	1.83	.51	.75
8E	4.01	1.73	.50	.75
8F	3.84	1.59	.39	.77
8G	3.80	1.55	.34	.77
8H	3.89	1.64	.39	.77

Como se pode observar no quadro 4, as correlações de cada item com o total (sem incluir o próprio item) apresentam o valor mínimo de .34 e o máximo de .66. O valor de alfa de Cronbach para esta escala foi o mais baixo encontrado, correspondendo ao valor de .78.

A estabilidade das pontuações foi avaliada num estudo separado com 116 participantes (104 do género feminino e 12 do género masculino, provenientes da população já mencionada, com a idade média de 19.65 anos e desvio-padrão igual a 3.46) que responderam duas vezes ao QFM com um intervalo que variou entre 45 e 70 dias (mediana do intervalo de reteste igual a 50 dias).

Os resultados totais obtidos apresentaram uma fiabilidade teste-reteste aceitável ($r_{xx} = .704$; $p < .01$).

Estudo da validade

Utilizou-se o método de análise em componentes principais com rotação *oblimin*. O critério clássico de Kaiser-Guttman, segundo o qual os factores a seleccionar devem ter um *eigenvalue* superior a 1, foi seguido em primeiro lugar. Também se recorreu ao *scree test* de Cattell tendo sido, então, sugerida a presença de quatro factores. A solução de quatro factores explica 35.2% da variância das respostas no QFM.

A matriz factorial é apresentada no quadro 5.

Quadro 5 - *Saturações de cada item na solução de quatro factores*

	F1	F2	F3	F4
Item 1	.54	.02	.16	.07
Item 2A	.38	-.01	.09	.09
Item 2B	.28	-.00	-.05	.03
Item 2C	.30	-.12	.23	.16
Item 2D	.43	-.03	.23	-.02
Item 2E	.41	-.01	.28	.05
Item 2F	.33	.09	.18	.09
Item 2G	.50	-.00	.20	.06
Item 2H	.45	-.08	.12	.06
Item 2I	.36	-.05	.26	.09
Item 2J	.38	.07	.22	.05
Item 2K	.49	.02	.21	.05
Item 2L	.39	.05	.29	.09
Item 2M	.40	.03	.20	.09
Item 2N	.28	-.08	.21	.14
Item 2O	.47	-.09	.31	.19
Item 2P	.57	-.01	.18	.07
Item 2Q	.61	-.05	.21	.09
Item 2R	.57	-.03	.20	.11
Item 3A	.54	-.01	.34	.03
Item 3B	.54	-.10	.51	-.00
Item 3C	.36	-.15	.71	-.01
Item 3D	.23	-.14	.78	-.02
Item 3E	.13	-.14	.74	.01
Item 4A	.40	-.05	.53	.23
Item 4B	.41	-.14	.70	.13
Item 4C	.29	-.19	.85	.10
Item 4D	.26	-.17	.88	.04
Item 4E	.17	-.14	.82	.06
Item 5A	.57	.04	.15	.16
Item 5B	.58	.01	.13	.11
Item 5C	.56	.02	.08	.00
Item 5D	.47	.01	.08	-.02
Item 6A	-.03	.46	-.06	-.05
Item 6B	-.00	.54	.03	-.04
Item 6C	.01	.63	-.10	.04
Item 6D	.16	.57	-.07	.32
Item 6E	-.02	.47	-.05	.21
Item 6F	.08	.53	-.12	.20
Item 6G	-.04	.48	-.15	.14
Item 6H	.02	.61	-.11	.04
Item 6I	.00	.70	-.08	.15
Item 6J	-.14	.52	-.10	.17
Item 6K	-.01	.59	-.08	.21
Item 6L	-.04	.62	-.12	.10
Item 6M	-.13	.64	-.16	.07
Item 6N	-.02	.55	-.12	.08

Item 6O	-.05	.78	-.14	.09
Item 6P	-.09	.77	-.08	.12
Item 6Q	.09	.77	-.06	.13
Item 6R	.04	.59	-.02	.07
Item 7A	.37	.04	-.09	.02
Item 7B	.45	.06	-.21	.07
Item 7C	.45	-.01	-.17	.06
Item 8A	.07	.03	.00	.61
Item 8B	.09	.13	.04	.81
Item 8C	.14	.05	.09	.80
Item 8D	.02	.06	.04	.66
Item 8E	-.03	.16	.10	.62
Item 8F	.14	.20	.04	.46
Item 8G	.03	.21	-.05	.44
Item 8H	.09	.12	-.01	.49
<i>Eigenvalues</i>	8.47	7.00	3.30	3.06
% da variância	13.67	11.28	5.31	4.94

Embora a solução factorial encontrada contemple 4 factores, como na versão original do QFM, estes não são exactamente sobreponíveis. Os factores designados de «Gravidade do esquecimento» e «Uso de mnemónicas», por Zelinski e colaboradores (1980), mantêm-se intactos (todos os itens 6 e todos os itens 8, respectivamente). Quanto aos factores originais «Frequência geral do esquecimento» (item 1, todos os itens 2, 3, 4 e 5) e «Funcionamento retrospectivo» (todos os itens 7) observaram-se novos arranjos dos itens.

Para nove itens (2A, 2B, 2C, 2F, 2I, 2J, 2L, 2N e 7A) não se obtiveram saturações iguais ou superiores a .40 em nenhum factor. Alguns destes itens (cujas saturações eram inferiores ou iguais a .30) foram eliminados (itens 2B, 2C e 2N). Verifica-se ainda que três itens (3B, 4A e 4B) apresentaram saturações iguais ou superiores a .40 em dois factores (F1 e F3).

Foram também calculados os valores de correlação entre as 4 escalas do QFM, que estão expressas no quadro 6.

Quadro 6 - Correlações entre as escalas do QFM

	Escala A	Escala B	Escala C	Escala D
Escala A	1.00			
Escala B	.45**	1.00		
Escala C	.02	-.14**	1.00	
Escala D	.17**	.02	.17**	1.00

** $p < .01$ (teste bilateral)

Legenda: Escala A = "Frequência geral do esquecimento e avaliação do funcionamento mnésico"; Escala B = "Frequência do esquecimento durante a leitura"; Escala C = "Gravidade do esquecimento" e Escala D = "Uso de mnemónicas".

Exceptuando o valor de correlação entre as escalas A e B, todos os restantes são baixos, o que poderá ser significativo que avaliam diferentes aspectos do funcionamento mnésico no quotidiano. Este resultado é concordante com o que é referido por Gilewski e colaboradores (1990).

Nenhum dos valores de correlação entre notas escolares (amostra com 151 participantes do 1º ano da licenciatura em psicologia, 136 do género feminino e 15 do género masculino; idade média de 19.36 anos e desvio padrão de 3.34) e as quatro escalas do QFM se revelou estatisticamente significativo.

Discussão e conclusões

A versão portuguesa do Questionário de Funcionamento da Memória apresentou propriedades psicométricas relativamente sólidas. Os resultados da análise da consistência interna com as quatro escalas revelaram-se elevados (acima de .84, com apenas uma excepção cujo valor foi de .78). O valor de fiabilidade teste-reteste para as pontuações totais foi de .70, sendo considerado aceitável para fins de investigação.

Quanto à validade estudada através do método de análise em componentes principais, encontrou-se uma solução com quatro factores, mas só dois são os mesmos da versão original do QFM: "Gravidade do esquecimento" e "Uso de mnemónicas". Os novos agrupamentos dos itens levaram à escolha, apenas indicativa, de outras designações, já mencionadas, para os restantes dois factores: "Frequência geral do esquecimento e avaliação do funcionamento mnésico" (item 1, todos os itens 2 – excepto 2B, 2C e 2N – item 3A, todos os itens 5 e todos os itens 7) e "Frequência do esquecimento durante a leitura" (todos os itens 3 – excepto 3A e todos os itens 4).

Os valores de correlação obtidos entre as várias escalas do QFM apontaram, de um modo geral, para uma avaliação diferenciada de aspectos da memória do quotidiano.

As correlações calculadas entre as quatro escalas do QFM e a classificação obtida numa disciplina do plano de estudos não se apresentaram como estatisticamente significativas.

Torna-se necessário recorrer a outros estudos de validação que incluam, por exemplo, a análise factorial confirmatória, a validade preditiva ou, ainda, a validade concorrente (o estudo com o Questionário de Metamemória na Idade Adulta de Dixon e Hultsch já está em curso). É também importante a utilização de amostras mais heterogéneas.

Por outro lado, importa igualmente considerar o estudo do efeito de variáveis que habitualmente constituem o foco de interesse no quadro da cognição social, i. e., de aspectos respeitantes a uma teoria dinâmica do *self*, que podem intervir durante o processo de recuperação de informação da memória. O conhecimento das *teorias implícitas* dos sujeitos, as suas próprias definições de memória e as variações individuais relativamente à *esquematicidade*, bem como a influência de estratégias ligadas à recuperação mnésica (p. ex., as chamadas heurísticas cognitivas: disponibilidade, acessibilidade e representatividade) na formação da resposta observável ou desempenho, deverá fazer parte dos estudos com questionários de metamemória.

Referências

- Ashcraft, M. H. (1989). *Human memory and cognition*. Glenview, IL: Scott, Foresman.
- Baddeley, A. (1999). *Human memory: Theory and practice* (revised edition). Hove: Psychology Press.
- Cavanaugh, J. C. e Baskind, D. (1996). Relations among basic processes, beliefs, and performance: A life-span perspective. In D. Herrmann, C. McEvoy, C. Hertzog, P. Hertel e M. K. Johnston (Eds.), *Basic and applied memory research: Practical applications* (Vol. 2, pp. 223-235). Mahwah, NJ: Erlbaum.
- Cavanaugh, J. C., Feldman, J. M. e Hertzog, C. (1998). Memory beliefs as social cognition: A reconceptualization of what memory questionnaires assess. *Review of General Psychology*, 2 (1), 48-65.
- Chow, S. L. (1987). Science, ecological validity and experimentation. *Journal for the Theory of Social Behavior*, 17 (2), 124-137.
- Gilewski, M. J., Zelinski, E. M. e Schaie, K. W. (1990). The Memory Functioning Questionnaire for assessment of memory complaints in adulthood and old age. *Psychology and Aging*, 5(4), 482-490.
- Herrmann, D. J. (1982). Know thy memory: The use of questionnaires to assess and study memory. *Psychological Bulletin*, 92 (2), 434-452.
- Hertzog, C., Dixon, R. A. e Hulstsch, D. F. (1990). Relationships between metamemory, memory predictions, and memory task performance in adults. *Psychology and Aging*, 5 (2), 215-227.
- Hertzog, C., Hulstsch, D. F. e Dixon, R. A. (1989). Evidence for the convergent validity of two self-report metamemory questionnaires. *Developmental Psychology*, 25(5), 687-700.
- Lane, C. J. e Zelinski, E. M. (2003). Longitudinal hierarchical linear models of the Memory Functioning Questionnaire. *Psychology and Aging*, 18 (1), 38-53.
- Markus, H. (1977). Self-schemata and processing information about the self. *Annual Review of Psychology*, 35, 299-337.
- McDonald-Miszczak, L., Hertzog, C. e Hulstsch, D. F. (1995). Stability and accuracy of metamemory in adulthood and aging: A longitudinal analysis. *Psychology and Aging*, 10 (4), 553-564.
- McDougall, G. J. (1998). Gender differences in coping and control with memory aging. *Journal of Women & Aging*, 10 (1), 21-40.
- Pinho, M. S. e Simões, M. R. (no prelo). Extensão de vocabulário técnico e desempenho académico num domínio específico de conhecimento. *Revista Portuguesa de Pedagogia*, 39.
- Pinto, A. C. (1997). Implicações escolares dos estudos de memória humana. *Psicologia, Educação e Cultura*, 1 (1), 31-50.

- Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review*, 96, 341-357.
- Simões, A. e Ferreira, J. A. (1996). Contributos para a avaliação da memória do dia-a-dia. *Actas das Jornadas de "Educação de Adultos em Portugal: Situação e Perspectivas"* (pp. 183-198). Coimbra: Núcleo de Assistência Psicológica e de Formação de Adultos da FPCE-UC.
- Zelinski, E. M., Gilewski, M. J. e Anthony-Bergstone, C. R. (1990). Memory Functioning Questionnaire: Concurrent validity with memory performance and self-reported memory failures. *Psychology and Aging*, 5(3), 388-399.
- Zelinski, E. M., Gilewski, M. J. e Thompson, L. W. (1980). Do laboratory tests relate to self-assessment of memory ability in the young and old?. In L. W. Poon, J. L. Fozard, L. S. Cermak, D. Arenberg e L. W. Thompson (Eds.), *New direction in memory and aging: Proceedings of The George A. Talland Conference* (pp. 519-544). Hillsdale, NJ: Erlbaum.

EVERYDAY MEMORY SELF-ASSESSMENT: SOME QUESTIONS CONCERNING THE USE OF METAMEMORY QUESTIONNAIRES AND STUDY OF PSYCHOMETRIC PROPERTIES OF THE MEMORY FUNCTIONING QUESTIONNAIRE (MFQ)

Maria Salomé Pinho

Mário R. Simões

Faculdade de Psicologia e de Ciências da Educação, Universidade de Coimbra

Abstract: The use of questionnaires to assess the self-appraisal of everyday memory functioning raises several concerns about the quality of measures among which was highlighted the lack of consistency and the “memory introspection paradox”. Some variables that influence response to questions involved in this type of questionnaires were pointed out within the social-cognition framework. Assuming that the knowledge about psychometric properties of memory self-ratings is an important issue we analysed the scores obtained in the Memory Functioning Questionnaire (MFQ) with a sample of 427 undergraduate students concerning reliability (internal consistency and test-re-test stability), construct validity (principal component analysis), internal validity (inter-correlations among MFQ scales), and correlations with school marks. In general, the Portuguese version of the MFQ showed good psychometric properties. However, the factorial solution found was unlike as for the group of items corresponding to two of the four original factors.

KEY-WORDS: *Memory, metamemory, questionnaires, reliability, validity.*

MOTIVAÇÃO INTRÍNSECA/EXTRÍNSECA DA VIVÊNCIA RELIGIOSA: UMA ESCALA REVISITADA

José H. Barros de Oliveira

Faculdade de Psicologia e de C. E., Universidade do Porto

Resumo

Objectivo principal deste estudo foi reanalisar a *Christian Religious Internalization Scale* (Escala de Interioridade Religiosa Cristã) de Ryan, Rigby e King (1993), já adaptada para a população portuguesa por Ferreira e Neto (2002). Agora alarga-se o seu estudo a outras amostras mais heterogéneas para melhor observar o seu comportamento estatístico. A amostra total é constituída por 387 sujeitos distribuídos por quatro subamostras: freiras, seminaristas, estudantes universitários e professores. Quer na análise estatística de cada amostra, quer na análise na sua globalidade, a escala apresenta dois factores, denominados no original como “identificação” e “introjecção”. Esta escala bifactorial manifestou possuir razoáveis propriedades psicométricas. Dado o seu aceitável funcionamento estatístico, avançamos ainda com uma análise de variância por sexo (não há diferenças significativas) e por grupos, manifestando os estudantes universitários um score mais baixo em “identificação” e as freiras um score mais elevado em “introjecção”. São discutidos os diversos resultados e feitas algumas sugestões a respeito da estrutura factorial da escala e ainda da terminologia usada para denominar os dois factores.

PALAVRAS-CHAVE: *Religião, avaliação, escala, interiorização, motivação.*

A religião é um fenómeno pessoal e social de grande relevância, que pode ser abordado desde as mais diversas perspectivas, incluindo as ciências humanas, como a psicologia. Einstein afirmava sem complexos: “sem a religião, a ciência é coxa; sem a ciência, a religião é cega”. Enquanto no estrangeiro, por exemplo, nos Estados Unidos (ultrapassados, quer o cientismo behaviorista quer os preconceitos psicanalíticos), encontramos milhares de trabalhos neste domínio, em Portugal, por razões que não importa agora aprofundar, mas que se prendem também com o nosso tradicional laicismo e agnosticismo, os psicólogos, salvo excepção (cf. Barros, 2000), desinteressaram-se por esta dimensão fundamental do ser humano. Todavia, há sinais de mudança também entre nós, como pode ser o número monográfico publicado

pela revista *Psicologia, Educação e Cultura* (2002, vol. 6, nº 2) sobre Psicologia da Religião.

Entre os múltiplos estudos sobre o fenómeno religioso, contam-se diversos Manuais de "Psicologia da Religião", como o de Hood et al. (1996). Outros autores investiram também na avaliação, na tentativa de compreender múltiplos aspectos da religião. Surgiram assim as mais diversas escalas, editadas num volume por Hill e Hood (1999). Entre tantos instrumentos de avaliação, conta-se a *Christian Religious Internalization Scale* (Escala de Interioridade Religiosa Cristã) de Ryan, Rigby e King (1993), já traduzida e adaptada para a população portuguesa por Ferreira e Neto (2002), numa amostra com estudantes universitários. Porém, pretendemos visitar ou reanalisar esta escala, alargando o seu estudo a outras amostras muito heterogéneas, que não apenas estudantes universitários, e sugerindo a supressão de alguns itens, a partir da análise factorial e também da análise de conteúdo dos itens. Propomos ainda o uso de uma terminologia mais adequada para designar cada um dos factores.

Efectivamente, Ferreira e Neto (2002) traduziram literalmente o inglês *internalization* por "internalização". Porém, o dicionário português não regista tal vocábulo, que pode ser considerado um anglicismo, sendo preferível traduzir por "interioridade" ou "interiorização", com a respectiva adjectivação: religião "interna", "interiorizada" ou "intrínseca", significa uma vivência da religião autodeterminada, motivada por razões pessoais livres ou assumida pelo indivíduo, verdadeiramente "interiorizada" (o que no original é denominado *identification*), versus uma religião "externa" ou "extrínseca", motivada por razões exteriores ao sujeito, como o julgamento dos outros ou a pressão social, ou então por razões pessoais mas compulsivas, como o medo ou a culpa (no original denominado *introjection*). Ferreira e Neto (2002) mantêm a tradução literal: "identificação" versus "introjecção".

Porém, estes termos pouco significam em português, se bem que poderíamos considerar uma religião intrínseca como "identificação", quando a pessoa se "identifica" com o que crê, e uma religião exterior ou imposta compulsivamente e/ou por conveniência social como "introjecção" à força. Todavia, a terminologia, essencialmente de origem psicanalítica, pode tornar-se confusa para expressar o que se pretende. Assim, propomos outra terminologia para nomear os dois factores: religião motivada intrinsecamente (religião intrínseca, assumida, livre) vs. religião motivada extrinsecamente (religião extrínseca, tradicional, compulsiva). Não obstante, podemos continuar a usar a terminologia original de Ryans et al. (1993), mantida por Ferreira e Neto (2002), desde que ela seja compreensível.

A “identificação” é entendida pelos psicanalistas como “processo psicológico pelo qual um indivíduo assimila um aspecto, uma propriedade, um atributo do outro e se transforma, total ou parcialmente, segundo o modelo dessa pessoa” (Laplace e Pontalis, 1979, s.v. identificação). Menos feliz ainda é o termo “introjecção” que na psicanálise é definido como “processo evidenciado pela investigação analítica: o indivíduo faz passar, de um modo fantasmático, de ‘fora’ para ‘dentro’, objectos e qualidades inerentes a esses objectos. A introjecção aproxima-se da incorporação, que constitui o seu protótipo corporal, mas não implica necessariamente uma referência ao limite corporal (introjecção no ego, no ideal do ego, etc.). (A introjecção) está estreitamente relacionada com a identificação” (*ibidem*, s.v. introjecção).

Fora do âmbito da psicanálise, os dois factores encontrados na escala em causa simplesmente fazem referência a uma religião assumida e vivida livremente pelo sujeito (interiorizada) ou não assumida, e por isso compulsiva ou imposta desde o exterior (não interiorizada). Teoricamente, no 2º factor (“introjecção”), ainda se deveria distinguir entre uma religião ‘imposta’ por constrangimentos externos ou ditame dos outros (itens 3,10,12) e uma religião ‘imposta’ por pressão ou compulsão do próprio sujeito (culpa ou medo – itens 5 e 7). De algum modo, também no factor dito “identificação” se poderia distinguir dois matizes: Uma religião vivida desde dentro (interiorizada) devido a causas verdadeiramente teológicas ou com pura motivação (praticamente só o item 1) ou devido a causas mais ou menos intimistas, gratificantes, ou menos ‘puras’ (itens 2, 4, 6, 8, 9,12).

Não obstante, pode prescindir-se destas ‘nuances’, porque é difícil uma ‘pureza’ total na vivência religiosa (partilhar a fé, amar a Deus, rezar, ir à igreja unicamente por amor a Deus, sem qualquer outro interesse ou motivação), como noutras manifestações ou emoções humanas, como o amor (mesmo uma mãe não ama o filho com total isenção de razões mais ou menos inconscientes que podem ter a ver com a autocompensação, a perspectiva de retribuição futura, etc.). Todavia, entre os místicos, é possível assistir a essa pureza total de intenções, expressa num amor incondicional, como quando um clássico espanhol (S. Francisco Xavier?) afirma que, mesmo que não houvesse inferno ou céu, amaria a Deus por Si mesmo, independentemente do medo do castigo ou da busca do prémio. O amor justifica-se por si mesmo, como afirma ainda S. Bernardo: “Amo porque amo, amo para amar; grande coisa é o amor”. Trata-se do amor “perfeito”, já expresso na Bíblia, como na 1ª carta de S. João: “No amor não há temor; pelo contrário, o perfeito amor lança fora o temor; de facto, o temor pressupõe castigo, e quem teme não é perfeito no amor” (4,18).

Postas estas considerações semânticas, psicológicas e mesmo teológicas, não despreciandas, que nos permitem compreender melhor o significado da escala de Ryan et al. (1993), não nos detemos em maior teoria sobre a interiorização da religião, teoria que pode ser lida na longa introdução que estes autores fazem à escala, aludindo a outras escalas similares, a começar pela de Allport e Ross (1967) sobre as diversas orientações religiosas, baseada na teoria de Allport (1950), um dos primeiros autores a estudar a religião do ponto de vista psicológico, distinguindo já entre orientação religiosa “intrínseca” (madura) e “extrínseca” (imatura ou utilitarista). O que realmente está em causa é a motivação da fé e da vivência religiosa (cristã). Muitos estudos, relacionando a religião com outras variáveis de personalidade, não são conclusivos ou são mesmo contraditórios por não distinguirem entre religião intrínseca e extrínseca.

Sendo o objectivo principal deste artigo estudar o comportamento psicométrico da escala, não avançamos com hipóteses, considerando o estudo exploratório.

Método

Participantes: A amostra global consta de 387 sujeitos, distribuídos por quatro subamostras: 1ª: 102 freiras de diversos Institutos Religiosos em Portugal (média de idade: 54 anos, entre os 25 e os 84 anos); 2ª: 77 estudantes de teologia de diversos Seminários Maiores Portugueses (média de idade: 21 anos, entre os 18 e os 28); 3ª: 110 estudantes da Faculdade de Ciências do Desporto e de Educação Física da Universidade do Porto (média de idade: 22 anos, entre os 18 e os 40; 58 eram rapazes e 52 raparigas; 4ª: 98 professores do ensino básico e secundário de diversos Colégios da região do Porto (média de idade: 35 anos, entre os 25 e os 60 anos; 46 homens e 52 mulheres).

Medidas: Os participantes respondiam inicialmente a alguns dados socio-demográficos. Era-lhes solicitado, a seguir, que preenchessem diversas escalas: 1) a *Escala de Interiorização religiosa cristã* (CRIS) de Ryan et al. (1993) com 12 itens e 4 modalidades de resposta, desde totalmente em desacordo até totalmente de acordo, já adaptada por Ferreira e Neto (2002); 2) um conjunto de breves escalas sobre *diversas perspectivas sobre a morte* de Spilka et al. (1977), com oito subescalas, a responder também em formato Likert, e que manifestaram possuir boas qualidades psicométricas na adaptação portu-
gue-

sa (Barros e Neto, 2004); 3) escala de *ansiedade face à morte* (Barros, 1998), em formato Likert, com suficientes propriedades psicométricas; 4) algumas escalas sobre diversas variáveis de personalidade, reflectindo emoções positivas e negativas, todas em formato Likert: *optimismo*, da autoria de Barros (1998), que manifestou boas qualidades psicométricas; *satisfação com a vida*, de Diener et al. (1985), adaptada por Neto, Barros e Barros (1990) para a população portuguesa, e que manifestou boas propriedades psicométricas; *felicidade*, da autoria de Barros (2001) e que se revelou capaz de avaliar o bem-estar subjectivo; *perdão* (Barros, 2002), com suficientes qualidades psicométricas; *neuroticismo*, que apresenta também razoáveis características psicométricas (Barros, 1999); *solidão*, de Russel et al. (1978), adaptada para a população portuguesa por Neto (1989, 1992), com um bom funcionamento psicométrico.

Procedimento: Este trabalho empírico foi realizado em 2002, em Portugal. Os questionários eram distribuídos aos interessados que os preenchiam no tempo mais aprazado, entregando posteriormente. Os alunos da Faculdade de Ciências do Desporto preencheram-nos nas aulas, na presença do professor.

Resultados e discussão

Visto tratar-se de populações muito heterogéneas e menos abordadas, nomeadamente o caso de freiras e seminaristas, entendeu-se por bem não se limitar à análise da amostra na sua globalidade, mas estudar antes o comportamento da escala em cada subamostra, funcionando cada uma como se de estudos diversos se tratasse ou de amostras independentes.

As características psicométricas da escala de Interiorização Religiosa Cristã foram verificadas, antes de mais, na 1ª amostra (freiras ou religiosas).

Para verificar a estrutura factorial da escala procedeu-se a uma análise factorial em componentes principais, aparecendo 4 componentes com eigenvalues superiores a 1, na matriz componencial, antes da rotação, embora bastante diferentes nos valores, sendo o 1º = 3.3, explicando só por si 27.2% da variância; o 2º = 2.7, explicando 22.4% da variância; o 3º = 1.2, explicando 10.3% e o 4º = 1.1, explicando 9.1, no total de 69% de variância explicada. Mas dado os dois primeiros factores serem mais fortes, explicando 50% da variância, e como no original inglês apenas apareceram dois factores superiores à unidade, foram rodados (rotação varimax) unicamente dois factores, cuja matriz componencial se apresenta no quadro 1:

Quadro 1: *Saturações dos itens nos dois componentes*

itens	1º componente	2º componente
1	.60	-.27
2	.42	.34
3	-.04	.78
4	.78	-.19
5	.03	.65
6	.81	.14
7	-.24	.77
8	.70	-.08
9	.66	.27
10	.12	.61
11	.61	.23
12	.16	.62

Como se pode verificar, arrumam-se claramente no 1º factor os itens 1,4,6,8,9,11 e menos claramente o 2, que satura relativamente baixo e nos dois factores, embora com tendência para o 1º componente ("identificação"), como aliás aconteceu no estudo de Ferreira e Neto (2002), ao contrário do estudo original dos autores da escala, na versão inglesa, onde o item 2 saturava predominantemente no 2º factor. Efectivamente, atendendo à análise de conteúdo do item, ele refere-se mais a uma religião interiorizada do que exteriorizada, embora a formulação seja de certo modo ambígua.

Quanto ao 2º factor ("introjecção") é representado pelos itens 3, 5, 7, 10 e 12.

A consistência interna das duas subescalas, avaliada pelo coeficiente alfa de Cronbach, foi de .78 para o 1º factor (com 7 itens, incluindo o 2) e .75 para o 2º factor (com 5 itens).

Pretendendo verificar a validade convergente e/ou discriminante da escala, procedeu-se a correlações entre cada um dos dois factores e outras escalas que avaliam construtos próximos relacionados com a morte (cf. Berman e Hays, 1999; Casebolt, 1999; Klug e Sinha, 1987; Maltby e Day, 2000; Osarchuk e Tatz, 1973; Ray e Najman, 1974, 1987; Schoenrade, 1989) (efectivamente, a religião e a morte andam de algum modo juntas, supondo-se até que o sentimento religioso em grande parte surgiu da tentativa de a pessoa se perpetuar no Além), ou outras dimensões teoricamente relacionadas com a vivência religiosa, quer pela positiva (optimismo, satisfação com a vida, felicidade, perdão), quer pela negativa (neuroticismo, solidão). No quadro 2 apresentam-se os respectivos valores (ao lado de r , consta a grandeza da significân-

cia – quando não existem diferenças significativas (n.s.) omite-se o valor da correlação).

Quadro 2: *Correlações entre os dois factores e as oito escalas breves de diversas perspectivas sobre a morte (a saber: morte como sofrimento e solidão; morte como uma vida do Além com recompensa; indiferença em relação à morte; morte como algo desconhecido; morte como abandono dos que dependem de nós; morte como coragem; morte como fracasso; morte como fim natural), a escala de ansiedade face à morte, outras escalas de variáveis de personalidade e ainda de um com o outro factor*

	Persp.1	persp. 2	persp. 3	persp. 4	persp. 5	persp. 6	persp. 7	persp. 8
Factor 1	n.s.	.n.s.	n.s.	n.s.	n.s.	.20*	n.s.	n.s.
Factor 2	.20*	.n.s.	n.s.		n.s.	.20*	.26**	.25**n.s
	Ans.morte	optimismo	satisfação	felicidade	perdão	neuroticismo	solidão	f. 2
Factor 1	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s	n.s.
Factor 2	.20*	n.s.	n.s.	n.s.	n.s.	n.s.	.26**	

* < .05; ** p < .01

Como se pode observar, o 1º factor correlaciona significativamente apenas com a morte vista como coragem. O 2º factor correlaciona com quatro perspectivas da morte e ainda com a escala de ansiedade face à morte, como também aconteceu no estudo de Ferreira e Neto (2002), levando a crer que uma religião mais “introjectada” esteja mais relacionada com o medo da morte. Ao contrário, não se verifica qualquer correlação significativa com as outras variáveis, a não ser com a solidão, no 2º factor, o que também aconteceu no estudo de Ferreira e Neto (2002), parecendo natural que uma vivência religiosa mais extrínseca possa levar a maior solidão. Também não há correlação entre os dois factores da escala em estudo.

Complementarmente, procedeu-se a uma análise de regressão, considerando todas as variáveis como independentes e a maior ou menor interiorização da religião como variável dependente. Em conformidade com as correlações, quanto ao 1º factor, a primeira e única variável a entrar significativamente na recta de regressão foi a morte, interpretada como coragem, com um Beta de .20. Quanto ao 2º factor, duas variáveis entraram significativamente na recta de regressão: morte como coragem e solidão, com pesos Beta, respectivamente de .26 e .25.

Quanto à 2ª amostra, com seminaristas (embora com número reduzido, na realidade trata-se de uma grande amostra, atendendo ao reduzido número da população total), eis as características psicométricas da escala de Interiorização Religiosa Cristã:

Na análise factorial, em componentes principais, apareceram igualmente 4 factores com *eigenvalues* superiores a 1, na matriz componencial, antes da rotação, embora bastante diferentes nos valores, sendo o 1º = 2.9, explicando só por si 24% da variância; o 2º = 2.2, explicando 18.3% da variância; o 3º = 1.3, explicando 11.1% e o 4º = 1.1, explicando 9.5, no total de 63% de variância explicada. Mas dado os dois primeiros factores serem mais fortes, explicando só por si 42% da variância e como no original inglês apenas apareceram dois factores superiores à unidade, foram rodados (rotação varimax) unicamente dois factores, cuja matriz componencial se apresenta no quadro 3:

Quadro 3: *Saturações dos itens nos dois componentes*

itens	1º componente	2º componente
1	.73	-.12
2	.66	.37
3	.10	.59
4	.73	-.14
5	.08	.70
6	.57	.20
7	-.15	.73
8	.65	-.32
9	.61	.01
10	-.22	.42
11	.28	-.12
12	-.09	.62

Como se pode verificar, arrumam-se no 1º factor ("identificação") os itens 1, 4, 6, 8, 9, e menos claramente o 2, que satura bem no 1º mas também relativamente bem no 2º factor. O item 11 nesta amostra não funciona. Quanto ao 2º factor ("introjecção") é representado pelos itens 3, 5, 7, 10 e 12, embora o 10 se mostre um pouco problemático.

A consistência interna das duas subescalas, foi de .72 para o 1º factor e .64 para o 2º, valores relativamente baixos.

Pretendendo verificar a validade discriminante, procedeu-se a correlações entre cada um dos factores e outras escalas que avaliam construtos próximos relacionados com a morte ou outras dimensões relativamente próximas, quer pela positiva (optimismo, satisfação com a vida, felicidade, perdão), quer pela negativa (neuroticismo, solidão). No quadro 4 apresentam-se os respectivos valores (ao lado do valor de *r*, consta a grandeza da significância):

Quadro 4: Correlações entre as duas subescalas e as oito pequenas escalas de perspectivas sobre a morte (a saber: morte como sofrimento e solidão; morte como uma vida do Além com recompensa; indiferença em relação à morte; morte como algo desconhecido; morte como abandono dos que dependem de nós; morte como coragem; morte como fracasso; morte como fim natural), a escala de ansiedade face à morte, outras escalas de variáveis de personalidade e ainda entre os dois factores

	Persp.1	persp. 2	persp. 3	persp. 4	persp. 5	persp. 6	persp. 7	persp. 8
Factor 1	n.s.	.n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s
Factor 2	n.s.	.n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s
	Ans.morte	optimismo	satisfação	felicidade	perdão	neuroticismo	solidão	f. 2
Factor 1	n.s.	.n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s
Factor 2	n.s.	.n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	

Como se pode observar, não se encontra nenhuma correlação significativa entre as diversas variáveis e a maior ou menor interiorização da religião. Também não há correlação entre os dois factores da escala em estudo. Consequentemente, não foi realizada a análise de regressão.

As características psicométricas da escala de Interiorização Religiosa Cristã foram verificadas, seguidamente, na 3ª amostra ou estudo com estudantes universitários.

Para verificar a estrutura factorial da escala, procedeu-se a uma análise factorial em componentes principais, aparecendo 2 factores com *eigenvalues* superiores a 1, na matriz componencial, antes da rotação, embora bastante diferentes nos valores, sendo o 1º = 4.7 explicando só por si 38.8% da variância, o 2º = 2.7, explicando 22.8% da variância, no total de 61.6% de variância explicada. Pcedeu-se seguidamente à rotação (varimax) destes dois componentes, cuja matriz componencial se apresenta no quadro 5:

Quadro 5: Saturações dos itens nos dois componentes

itens	1º componente	2º componente
1	.73	.26
2	.52	.52
3	.21	.73
4	.85	.16
5	.38	.57
6	.90	.05
7	.10	.79
8	.85	.15
9	.85	-
10	.48	.39
11	.64	.24
12	-.10	.81

Vão claramente no 1º factor ("identificação") os itens 1,4,6,8,11, saturando o 2 do mesmo modo nos dois factores, notando-se a ambiguidade do item novamente nesta amostra. Quanto ao 2º factor ("introjecção") é representado claramente pelos itens 3, 7, e 12 e menos claramente pelo item 5, que satura também bastante no 1º factor. O item 10 também se comporta ambiguamente, saturando nos dois factores.

A consistência interna das duas subescalas, foi de .90 para o 1º factor e de .75 para o 2º.

Pretendendo verificar a validade discriminante da escala, procedeu-se a correlações entre cada uma destas duas subescalas e outras escalas que avaliam construtos mais ou menos relacionados. No quadro 6 apresentam-se os respectivos valores (ao lado do valor de r , consta a grandeza da significância):

Quadro 6: *Correlações entre as duas subescalas e as oito pequenas escalas de perspectivas sobre a morte (a saber: morte como sofrimento e solidão; morte como uma vida do Além com recompensa; indiferença em relação à morte; morte como algo desconhecido; morte como abandono dos que dependem de nós; morte como coragem; morte como fracasso; morte como fim natural), a escala de ansiedade face à morte, outras escalas de variáveis de personalidade e entre os dois factores:*

	Persp.1	persp. 2	persp. 3	persp. 4	persp. 5	persp. 6	persp. 7	persp. 8
Factor 1	n.s.	.n.s.	n.s.	n.s.	n.s.	.21*	n.s.	n.s.
Factor 2	n.s.	.30**	n.s.	n.s.	n.s.	.23**	n.s.	-.33***
	Ans.morte	optimismo	satisfação	felicidade	perdão	neuroticismo	solidão	f. 2
Factor 1	n.s.	n.s.	n.s.	n.s.	.23**	n.s.	n.s.	.50***
Factor 2	n.s.	n.s.	n.s.	-.22**	n.s.	.28**	.36***	

* $p < .05$; ** $p < .01$; *** $p < .001$

Quanto ao 1º factor, correlaciona significativamente apenas com a morte vista como coragem e ainda com a escala do perdão. O 2º factor correlaciona significativamente com três perspectivas da morte e com diversas escalas: da felicidade (pela negativa), do neuroticismo e da solidão. Verifica-se ainda uma alta correlação entre os dois factores da escala em estudo, sinal de que os estudantes universitários relacionam bastante ou não vêem grande diferença entre a "identificação" e a "introjecção", ao contrário do que aconteceu com as duas amostras anteriores com pessoas religiosas.

Complementarmente, procedeu-se a uma análise de regressão, considerando todas as variáveis como independentes e a maior ou menor interiorização da religião como variável dependente. Em conformidade com as correlações, quanto ao 1º factor, a primeira e única variável a entrar significativamente na recta de regressão foi o perdão, com um Beta de .22. Quanto ao 2º factor, várias variáveis entraram significativamente na recta de regres-

são, pela seguinte ordem: solidão, morte como crença no Além, neuroticismo e morte como coragem, com Betas de .35, .28, .24, .20.

Procedeu-se ainda a uma análise de variância para ver se havia diferenças por sexo (58 M; 52 F). Não foram encontradas diferenças significativas, quer na subescala de "identificação", quer na de "introjecção". Também os autores da escala não encontraram diferenças. Ferreira e Neto (2002) obtiveram diferenças significativas na subescala de "identificação", a favor das raparigas; na outra subescala não houve diferenças por sexo.

As características psicométricas da escala de Interiorização Religiosa Cristã foi verificada, seguidamente na 4ª amostra ou estudo com professores.

Para verificar a estrutura psicométrica da escala, procedeu-se a uma análise factorial em componentes principais, aparecendo 2 factores com *eigenvalues* superiores a 1, na matriz componencial, antes da rotação, embora bastante diferentes nos valores, sendo o 5º = 5.0, explicando só por si 42% da variância, o 2º = 2.0, explicando 16.5% da variância, no total de 58% de variância explicada. Pcedeu-se seguidamente à rotação (varimax) destes dois componentes, cuja matriz componencial se apresenta no quadro 7:

Quadro 7: *Saturações dos itens nos dois componentes*

itens	1º componente	2º componente
1	.77	.30
2	.68	.50
3	.01	.68
4	.80	.12
5	.27	.56
6	.88	.01
7	.02	.74
8	.83	.05
9	.89	.03
10	.85	.31
11	.62	.43
12	-.05	.75

Saturam claramente no 1º factor ("identificação") os itens 1,4,6,8,9,10; este último, nas amostras anteriores, saturava predominantemente no 2º factor. O item 2 satura do mesmo modo nos dois factores, notando-se novamente a ambiguidade do item. O item 11 também satura bem nos dois factores, comportando-se nesta amostra ambigualmente, como já tinha acontecido noutras. Quanto ao 2º factor ("introjecção") é representado claramente pelos itens 3, 7, e 12 e menos claramente pelo item 5, que satura também razoavelmente

no 1º factor. Por conseguinte, a estrutura factorial nesta amostra mostra-se menos clara.

A consistência interna das duas subescalas, foi de .91 para o 1º factor e de .68 para o 2º.

Pretendendo verificar a validade discriminante da escala, procedeu-se a correlações entre cada um destes dois factores e outras escalas que avaliam construtos mais ou menos relacionados. No quadro 8 apresentam-se os respectivos valores.

Quadro 8: *Correlações entre as duas subescalas e as oito pequenas escalas de perspectivas sobre a morte (a saber: morte como sofrimento e solidão; morte como uma vida do Além com recompensa; indiferença em relação à morte; morte como algo desconhecido; morte como abandono dos que dependem de nós; morte como coragem; morte como fracasso; morte como fim natural), a escala de ansiedade face à morte, outras escalas de variáveis de personalidade e ainda entre os dois factores*

	Persp.1	persp. 2	persp. 3	persp. 4	persp. 5	persp. 6	persp. 7	persp. 8
Factor 1	n.s.	.50***	n.s.	n.s.	n.s.	.22*	n.s.	n.s.
Factor 2	.31**	.29**	n.s.	n.s.	.34**	n.s.	.32**	n.s.
	Ans.morte	optimismo	satisfação	felicidade	perdão	neuroticismo	solidão	f. 2
Factor 1	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	.44***
Factor 2	.28**	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	

* $p < .05$; ** $p < .01$; *** $p < .001$

Quanto ao 1º factor, correlaciona significativamente unicamente com a morte considerada como crença no Além (altamente significativo) e como coragem. No 2º factor, como sempre, aparecem mais correlações significativas: a morte como sofrimento e solidão, como crença na vida do Além, como abandono dos que dependem de nós, como fracasso, e ainda com a escala de ansiedade face à morte. Verifica-se também uma alta correlação entre os dois factores da escala, sinal de que, neste estudo, os dois factores não se manifestam muito diferentes.

Complementarmente, procedeu-se também a uma análise de regressão, considerando todas as variáveis como independentes e a maior ou menor interiorização da religião como variável dependente. Em conformidade com as correlações, quanto ao 1º factor, a primeira e única variável a entrar significativamente na recta de regressão foi a morte como crença no Além, com um Beta de .50. Quanto ao 2º factor, três variáveis entraram significativamente na recta de regressão, pela seguinte ordem: morte como abandono dos que dependem de nós, morte como crença no Além e morte como fracasso, com Betas respectivamente de .34, .25, .22.

Procedeu-se ainda a uma análise de variância para ver se havia diferenças por sexo (46 M; 52 F). Não foram encontradas diferenças significativas,

quer na subescala de “identificação”, quer na de “introjecção”, como aconteceu na amostra anterior.

Depois de estudada cada uma das subamostras, pareceu útil também analisar a amostra no seu conjunto, apesar da sua heterogeneidade, mas para realçar melhor a estrutura da escala e confirmar algumas anomalias já verificadas singularmente. Assim, as características psicométricas da escala de Interiorização Religiosa Cristã foram verificadas na *amostra global*.

Procedeu-se, antes de mais, a uma análise factorial em componentes principais, apresentando-se 2 factores com *eigenvalues* superiores a 1, na matriz componencial, antes da rotação, embora muito diferentes nos valores, sendo o 1º = 5.3 explicando só por si 44.4% da variância e o 2º = 1.9, explicando 16.4% da variância, explicando no conjunto 60.8 da variância. Já Ferreira e Neto (2002) tinham encontrado um 1º componente muito mais forte que o 2º, o mesmo acontecendo no estudo original de Ryan et al. (1993). Isto está a indicar a tendência para um único factor, ou ao menos a ‘fraqueza’ do 2º em relação ao 1º que de qualquer forma já inclui, pela negativa, o 2º. Todavia, guiados pelos resultados de Ryan et al. (1993) foram rodados os dois factores, cujas saturações se apresentam no quadro 9:

Quadro 9: *Saturações dos itens nos dois componentes:*

itens	1º componente	2º componente
1	.83	.18
2	.72	.36
3	.19	.72
4	.87	.18
5	.40	.59
6	.83	-
7	-	.80
8	.86	-
9	.83	-
10	.18	.52
11	.68	.22
12	-	.74

Como se pode verificar, arrumam-se claramente no 1º factor os itens 1,4,6,8,9,11 e menos nitidamente o 2, embora seja grande a tendência para o 1º componente (“identificação” - religião intrínseca), como aliás aconteceu no estudo de Ferreira e Neto (2002), ao contrário do estudo original dos autores da escala, na versão inglesa, onde o item 2 saturava predominantemente no 2º factor. Efectivamente, atendendo à análise de conteúdo do item, ele refe-

re-se mais a uma religião interiorizada do que exteriorizada, embora a formulação esteja ferida de certa ambiguidade.

Quanto ao 2º factor ("introjecção" - religião extrínseca) é representado claramente pelos itens 3, 7, 10, 12 e menos claramente pelo item 5 (que satura também bastante fortemente no 1º factor). Em conclusão, poderíamos excluir o item 2 no 1º factor e o 5 no 2º. Porém, para não reduzir mais o número de itens em cada um dos factores, eles podem ser mantidos, particularmente o 5.

A consistência interna das duas subescalas, avaliada pelo coeficiente alfa de Cronbach, foi de .92 para o factor de "identificação" e .73 para o factor de "introjecção", sendo mais elevado no 1º, como acontece no estudo português de Ferreira e Neto (2002) (nos estudos com a escala original em inglês, numa das amostras os alfas eram idênticos nos dois factores, noutra era mais alto na identificação e noutra na introjecção; notar que no original o 2º factor conta com mais um item (o nº 2) e por isso o 1º factor com menos um). Todavia, parece evidente que o 1º factor é mais consistente, atendendo à análise factorial e à fidelidade.

Pretendendo verificar a validade convergente e discriminante, procedeu-se a correlações entre cada uma destas subescalas e outras escalas que avaliavam construtos próximos relacionados com a morte (efectivamente a religião e a morte andam juntas, supondo-se até que o sentimento religioso em grande parte surgiu da tentativa de a pessoa se perpetuar no Além) ou outras dimensões relativamente próximas, quer pela positiva, se se trata duma vivência religiosa verdadeira (optimismo, satisfação com a vida, felicidade, perdão), quer pela negativa (neuroticismo, solidão). No quadro 10 apresentam-se os respectivos valores (ao lado do valor de r , consta a grandeza da significância; os valores não significativos são expressos como n.s.)

Quadro 9: *Correlações entre as duas subescalas e as oito pequenas escalas de perspectivas sobre a morte (a saber: morte como sofrimento e solidão; morte como uma vida do Além com recompensa; indiferença em relação à morte; morte como algo desconhecido; morte como abandono dos que dependem de nós; morte como coragem; morte como fracasso; morte como fim natural), a escala de ansiedade face à morte, outras escalas de variáveis de personalidade e ainda entre os dois factores*

	Persp.1	persp. 2	persp. 3	persp. 4	persp. 5	persp. 6	persp. 7	persp. 8
Factor 1	n.s.	.40***	n.s.	n.s.	n.s.	.23***	n.s.	n.s.
Factor 2	.16**	.36***	n.s.	n.s.	n.s.	.25***	n.s.	n.s.
	Ans.morte	optimismo	satisfação	felicidade	perdão	neuroticismo	solidão	f. 2
Factor 1	n.s.	n.s.	n.s.	n.s.	.19***	n.s.	n.s.	.35***
Factor 2	n.s.	n.s.	n.s.	-.17**	.14**	.17**	.26***	

** $p < .01$; *** $p < .001$

Como se pode observar, quer o 1º, quer o 2º, correlacionam significativamente com a morte vista como crença no Além e como coragem. O 2º factor correlaciona também, embora de forma menos significativa, com a morte interpretada como sofrimento e solidão. Nota-se ainda uma correlação significativa entre o 1º factor e o perdão e entre o 2º factor e a felicidade (negativamente), o perdão, o neuroticismo e, mais significativamente, a solidão, como aconteceu no estudo de Ferreira e Neto (2002).

Algumas destas correlações são difíceis de explicar, mas pode conceber-se que uma religião mais ou menos interiorizada pode ter a ver com a crença na vida eterna e com um afrontamento corajoso da morte. Pode também compreender-se a correlação com o perdão, supondo que a religião mais ou menos interiorizada possa predispor melhor a pessoa para o perdão (cf. Barros, 2000, 2002). A correlação entre a vivência de uma religião extrínseca e o neuroticismo e a solidão também parecem ter uma certa lógica, bem como a correlação negativa com a felicidade (cf. Barros 2001).

Nota-se ainda uma correlação bastante significativa entre os dois factores, sinal de que os dois factores apresentam certas semelhanças, embora sejam distintos.

Complementarmente, procedeu-se também a uma análise de regressão, considerando todas as variáveis como independentes e a maior ou menor interiorização da religião como variável dependente. Em conformidade com as correlações, quanto ao 1º factor, a primeira variável a entrar significativamente na recta de regressão foi a morte como crença na vida do Além, com um Beta de .40, seguindo-se a morte como coragem, com um Beta de .25. Quanto ao 2º factor, três variáveis entraram significativamente na recta de regressão: morte como crença no Além, solidão, morte como coragem, com pesos Beta, respectivamente de .36, .24 e .17.

Procedeu-se, finalmente, a uma análise de variância (*oneway*) para verificar se havia diferenças significativas entre os quatro grupos quanto à vivência da religião mais ou menos assumida. Quanto ao 1º factor (religião motivada intrinsecamente - 'identificação') verificaram-se diferenças altamente significativas ($F(3/383)=11.9$; $p<.001$) entre o 3º grupo (estudantes universitários) e os restantes grupos (freiras, seminaristas, professores), apresentando os estudantes uma média mais baixa (15.7) em relação aos 1º, 2º e 4º grupos (respectivamente, com médias de 17.6, 18.9, 18.0 - quanto maior a média, maior "identificação"), sinal de uma vivência religiosa menos motivada intrinsecamente por parte dos alunos universitários. Isto parece conforme com a bibliografia que normalmente atribui aos mais novos menor motivação intrínseca religiosa (cf. Barros, 2000; Hood et al., 1996). Os estudantes seminaristas

têm a mesma idade, mas trata-se de profissionais do religioso com ‘obrigação’ de viverem a religião mais convictamente.

No que concerne ao 2º factor (religião motivada extrinsecamente - ‘introjecção’) notam-se diferenças significativas ($F(3/383)=16.7$; $p<.001$) entre o 1º grupo (religiosas), apresentando uma média de 10.7, e os restantes grupos com médias, respectivamente de 9.4, 8.1, 8.7 (quanto maior a média, maior “introjecção”), sendo difícil explicar como é que as freiras vivem a religião com motivações mais extrínsecas, quer de ordem pessoal quer social, seguindo-se outro grupo também religioso (seminaristas). Talvez a educação religiosa nos conventos, feita mais pelo temor do que pelo amor, e ainda um certo escrúpulo, possa de algum modo explicar esta situação.

Conclusão

O objectivo principal deste estudo foi observar o comportamento psicométrico da escala de interiorização religiosa, confrontando os resultados das diversas amostras, mais heterogêneas no tipo de pessoas e na idade do que a de Ferreira e Neto (2002), que usaram unicamente estudantes universitários, podendo assim verificar-se também a validade externa da escala. Através da análise factorial, confirmada pela análise de conteúdo dos itens, foram identificados dois factores, manifestando-se o 1º mais forte, mesmo porque inclui implicitamente o 2º (quanto maior for a vivência interiorizada da religião, mais baixa será a sua prática compulsiva, por culpa, medo ou pressão social).

Apesar da brevidade da escala, considerada bifactorial, ela manifestou suficiente consistência interna (fidelidade). Dado os resultados do nosso estudo coincidirem fundamentalmente com os de Ferreira e Neto (2002), pode considerar-se esta escala fidedigna e válida para avaliar o grau de interiorização da religião. Não obstante, algumas observações e correcções devem ser feitas.

Em primeiro lugar, atendendo à análise factorial em cada uma das amostras, pode excluir-se, em futuros estudos, o item 2 que satura tendencialmente melhor no 1º factor (“identificação”), a exemplo do que aconteceu no estudo de Ferreira e Neto (2002), ao contrário do original de Ryans et al (1993), mas também satura bem ou mesmo igual no 2º factor. Dada esta ambiguidade, que se nota outrossim numa análise de conteúdo do item, votamos pela sua exclusão.

Num segundo tempo, e para termos uma escala mais apurada, poderíamos também excluir os itens 10 (nas duas amostras com religiosos satura prin-

principalmente no 2º factor, mas nas outras duas com leigos, satura predominantemente no 1º) e 11 (na amostra com seminaristas não funciona e noutras duas satura também bastante no 2º factor). Por seu lado, o item 5, satura bem no 2º factor em duas amostras, com pessoas religiosas, mas menos bem nas outras duas, mostrando-se algo problemático; todavia, pode ser mantido. Outrossim o item 9 é praticamente uma redundância ou repetição do item 8, mas, como funciona bem, pode ser mantido.

Assim, teríamos a representar o 1º factor os itens 1,4,6,8,9 e o 2º os itens 3,5,7,12 manifestando estes uma religião menos interiorizada, motivada não intrinsecamente mas por pressão social (itens 3 e 12), ou pelo sentimento de culpa (item 5) ou por medo de Deus (item 7). Se quiséssemos ser mais radicais, poderíamos optar por uma escala unifactorial, usando apenas os cinco itens indicativos do 1º factor (uma religião vivida, assumida conscientemente, bem motivada), que seria manifestada consoante os valores médios fossem mais elevados; ao contrário, médias baixas estariam a indicar uma religião menos consciente, mais compulsiva ou menos livre. Todavia, como este 2º factor indicativo de uma religião menos motivada intrinsecamente pode ser devido à pressão social ou então por sentimento de culpa ou por medo de Deus, haverá conveniência em manter a escala bifactorial.

Poderia até pensar-se numa escala tetrafactorial, distinguindo, no 2º factor ("introjecção"), entre religião extrínseca por razões mais pessoais (culpa ou medo de Deus) ou por razões sociais (pressão externa). Mesmo no 1º factor ("identificação") também se poderia distinguir entre motivações 'puras' (amor incondicional a Deus) e motivações mais ou menos "gratificantes" (como se expressam dois itens) onde pode estar presente certo egoísmo ou narcisismo. De facto, nas duas primeiras amostras, com pessoas religiosas, que têm obrigação de captar melhor estas 'nuances', apareceram, na análise factorial, quatro valores próprios superiores à unidade, ao contrário das outras amostras onde só apareceram dois. Todavia, pensar numa escala deste género, com quatro factores, seria esmiuçar demasiado as motivações religiosas, tornando-se ao mesmo tempo necessário introduzir novos itens na escala.

De qualquer modo, o 2º factor é sempre mais problemático (aliás isso acontece noutras escalas), como se pode também verificar pelos coeficientes alfa, sendo sempre maior a consistência interna do 1º factor. Acontece ainda que nas duas amostras com leigos a correlação entre os dois factores é bastante elevada, ao contrário do que sucede nas amostras com freiras e seminaristas, sinal de que esta população religiosa sabe distinguir melhor entre uma religião mais ou menos bem assumida e motivada.

No que concerne às correlações, devemos assumir que algumas variáveis praticamente nada têm a ver com o grau de interiorização da religião, pois em nenhuma das amostras se apresentaram correlações significativas. Quanto às oito pequenas escalas sobre diversas perspectivas sobre a morte, adaptadas por Barros e Neto (2004), supondo-se a priori e por estudos empíricos que há correlação entre a religião e diversas atitudes ou perspectivas frente à morte (cf. Berman e Hays, 1999; Casebolt, 1999; Klug e Sinha, 1987; Maltby e Day, 2000; Osarchuk e Tatz, 1973; Ray e Najman, 1974, 1987; Schoenrade, 1989), na realidade só se encontraram correlações significativas entre a religião interiorizada (1º factor) e a morte interpretada como coragem, o que é compreensível, supondo-se que a pessoa religiosa pode aceitar a morte ou morrer com maior tranquilidade. Na amostra com professores também apareceu uma alta correlação com a morte interpretada como recompensa na vida do Além, o que também parece evidente; todavia, nas outras três amostras não se manifestou esta correlação. Ainda tendo em conta unicamente o 1º factor, não se encontrou qualquer correlação entre a motivação religiosa e as outras seis escalas sobre diversas variáveis da personalidade, a não ser na amostra dos estudantes onde aparece uma correlação significativa com o perdão, supondo-se que a pessoa verdadeiramente religiosa está mais disposta a perdoar (cf. Barros, 2004).

Atendendo ao 2º factor (religião “introjectada”) encontram-se mais correlações significativas com a morte, entendida, como no 1º factor, como coragem e como vida do Além, mas ainda como sofrimento e solidão, como abandono dos que dependem de nós, como fracasso e como fim natural (neste caso a correlação é negativa). Em duas amostras, aparecem também correlações significativas com a escala de ansiedade face à morte, como acontece no estudo de Ferreira e Neto (2002). Parece, assim, que as pessoas com uma vivência religiosa menos motivada e esclarecida se encontram mais confrontadas e com mais medo da morte.

Quanto às outras escalas sobre diversas variáveis de personalidade, as correlações mais elevadas foram com a solidão, em duas amostras, encontrando-se ainda na amostra com estudantes uma correlação significativa com o neuroticismo e com a felicidade (neste caso, correlação negativa), estando a significar que a pessoa com religiosidade mais ou menos extrínseca pode sentir mais a solidão (também Ferreira e Neto, 2002, encontraram uma correlação significativa entre a religião “introjectada” e a solidão e com a satisfação com a vida (cor. negativa) (cf. ainda Neto e Barros, 2002). Quanto ao neuroticismo, também Ryan et al. (1993), no seu estudo original da escala sobre a interiorização da religião tinham concluído que a religião verdadeiramente in-

teriorizada (“identificação”) levava a uma saúde mental melhor, ao contrário da religião “introjectada”.

Também Martins (2002) conclui, num estudo interessante, que a abordagem do problema religioso e espiritual em psicoterapia, facilitaria a cura, apesar de, quer psicoterapeutas, quer clientes, constatarem que isso não acontece. Já outros autores tinham estudado a correlação entre valores religiosos, a psicoterapia e a saúde mental (Bergin, 1980, 1983), tendo Ellis (1980), em reacção ao primeiro artigo de Bergin (1980), defendido os “valores ateus” em psicoterapia, certamente com complexos, na linha freudiana, de que a religião denotaria patologia. Actualmente, os psicólogos e psicoterapeutas humanistas e sistémicos pensam que é de respeitar e promover os valores religiosos como coadjuvantes da saúde mental.

Voltando à escala em análise, dado que ela manifestou propriedades psicométricas razoáveis, procedemos também a uma análise de variância por grupos e por sexo. No que concerne aos quatro grupos ou amostras, foram encontradas diferenças significativas, quanto ao 1º factor, unicamente entre o grupo de estudantes universitários e os restantes, supondo-se que os mais novos não interiorizam tão bem a religião (cf. Barros, 2000; Hood et al., 1996). No que tange ao 2º factor, assiste-se a diferenças significativas entre o 1º grupo (de freiras) e os restantes grupos, estando a indicar uma tendência maior nas religiosas a viver uma religião mais extrínseca ou menos interiorizada, fruto talvez da educação mais restrita ou de um maior rigor interpretativo dos itens. Outra análise de variância por sexo, nos dois últimos grupos, não deu diferenças significativas, na sequência de outros estudos (Ferreira e Neto, 2002; Ryan et al., 1993).

Em futuros estudos pode usar-se a escala reduzida, omitindo os itens mencionados, ou então passar apenas os itens do 1º factor, considerando a escala unifactorial. Em qualquer dos casos, é necessário reestudar de novo o comportamento estatístico desta eventual redução da escala. Para avaliar a sua validade convergente e/ou discriminante é necessário também passar conjuntamente algumas escalas no âmbito da religião, como a escala de atitudes frente à religião (cristã) de Francis e Stubbs (1987), estudada numa amostra portuguesa por Ferreira e Neto (2002), mas publicada em revista estrangeira.

Referências

- Allport, G. (1950). *The individual and his religion*. New York: MacMillan.
- Allport, G. e Ross, J. (1967). Personal religious orientation and prejudice. *Journal of Personality and Social Psychology*, 5, 432-443.
- Barros, J. (1998). *Viver a morte – Abordagem antropológica e psicológica*. Coimbra: Almedina.
- Barros, J. (1998). Optimismo: Teoria e avaliação (proposta de uma nova escala). *Psicologia, Educação e Cultura*, 2 (2), 295-308.
- Barros, J. (1999). Neuroticismo: teoria e avaliação (proposta de uma nova escala). *Psicologia, Educação e Cultura*, 3 (1), 129-144.
- Barros, J. (2000). *Psicologia da Religião*. Coimbra: Almedina.
- Barros, J. (2001). Felicidade: Natureza e avaliação (proposta de uma nova escala). *Psicologia, Educação e Cultura*, 5 (2), 289-318.
- Barros, J. (2002). Perdão: Teoria e avaliação (proposta de uma nova escala). *Psicologia, Educação e Cultura*, 6 (2), 303-320.
- Barros, J. (2004). *Psicologia Positiva*. Porto: ASA.
- Barros, J. e Neto, F. (2004). Validação de um instrumento sobre diversas perspectivas da morte. *Análise Psicológica*, 22 (2), 355-367.
- Bergin, A. (1980). Psychotherapy and religious values. *Journal of Consulting and Clinical Psychology*, 48, 95-105.
- Bergin, A. (1983). Religiosity and mental health: A critical reevaluation and meta-analysis. *Professional Psychology: Research and Practice*, 14 (2), 170-184.
- Berman, A. e Hays, J. (1973). Relation between death anxiety, belief in afterlife, and locus of control. *Journal of Consulting and Clinical Psychology*, 41 (2), 318.
- Casebolt, J. (1999). Scales of Death Afterlife, Wiews of. In P. Hill e R. Hood (Ed.), *Measures of Religiosity* (pp. 431-445). Birmingham, AL: Religious Education Press.
- Diener, E., Emmons, R., Larsen, R., e Griffins, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, 49 (1), 71-75.
- Ellis, A. (1980). Psychotherapy and atheistic values: A response to A. E. Bergin's "Psychotherapy and religious values". *Journal of Consulting and Clinical Psychology*, 48, 635-639.
- Ferreira, A. e Neto, F. (2002). Dois tipos de internalização religiosa: Introjecção e identificação. *Psicologia, Educação e Cultura*, 6 (2), 321-334.
- Ferreira, A. e Neto, F. (2002). Psychometric properties of the Francis Scale of Attitude towards Christianity among portuguese university students. *Psychological Reports*, 91, 995-998.

- Francis, L. e Stubbs, M. (1987). Measuring attitudes towards Christianity: from childhood into adulthood. *Personality and Individual Differences*, 8, 741-743.
- Hill, P. e Hood, R. (Ed.) (1999). *Measures of Religiosity*. Birmingham, Al: Religious Education Press.
- Hood, R., Spilka, B., Hunsberger, B., e Gorsuch, R. (1996). *The Psychology of Religion*. New York: The Guildford Press.
- Klug, L. e Sinha, A. (1987). Death acceptance: A two-component formulation and scale. *Omega*, 18 (3), 229-235.
- Laplanche, J. e Pontalis, J.-B. (1979). *Vocabulário da Psicanálise* (5ª edição). Lisboa: Moraes Editores.
- Maltby, J. e Day, L. (2000). Religious orientation and death obsession. *The Journal of Genetic Psychology*, 161 (1), 122-124.
- Martins, J (2002). Vivências sobrenaturais e psicoterapia. *Psicologia, Educação e Cultura*, 6 (2), 335-348.
- Neto, F. (1989). A escala de solidão da UCLA: Adaptação portuguesa. *Psicologia Clínica*, 2, 65-79.
- Neto, F. (1992). *Solidão, embaraço e amor*. Porto: Centro de Psicologia Social.
- Neto, F., Barros, J., e Barros, A. (1990). Satisfação com a vida. In L. Almeida, et al. (Eds), *A ação educativa: análise psicossocial* (pp. 105-117). Leiria: ESEL/APPORT.
- Neto, F. e Barros, J. (2002). Satisfação com a vida: Diferenças conforme a religião e a cultura. *Psicologia, Educação e Cultura*, 6 (2), 349-361.
- Osarchuk, M. e Tatz, S. (1973). Effect of induced fear of death on belief in afterlife. *Journal of Personality and Social Psychology*, 27 (2), 256-260.
- Ray, J. e Najman, J. (1974). Death anxiety and death acceptance: A preliminary approach. *Omega*, 5 (4), 311-315.
- Ray, J. e Najman, J. (1987). Neoconservatism, mental health and attitude to death. *Personality and Individual Differences*, 8 (2), 277-279.
- Russel, D., Perpalu, L., e Ferguson, M. (1978). Developing a measure of loneliness. *Journal of Personality Assessment*, 42, 290-294.
- Ryan, R., Rigby, S. e King, K. (1993). Two types of religious internalization and their relations to religious orientations and mental health. *Journal of Personality and Social Psychology*, 65 (3), 586-596.
- Spilka, B., Stout, L., Minton, B., e Sizemore, D. (1977). Death and personal faith: A psychometric investigation. *Journal for the Scientific Study of Religion*, 16, 169-178.
- Schoenrade, P. (1989). When I die...Belief in afterlife as a response to mortality. *Personality and Social Psychology Bulletin*, 15 (1), 91-100.

THE INTRINSIC AND EXTRINSIC MOTIVATION OF RELIGIOUS EXPERIENCE: A REVISITED SCALE

José H. Barros de Oliveira

Faculdade de Psicologia e de C. E., Universidade do Porto, Portugal

Abstract: The main goal of this study was to re-analyze the Ryan, Rugby and King (1993) *Christian Religious Internalization Scale*, adapted to the Portuguese population by Ferreira and Neto (2002). The scale is here expanded to cover other heterogeneous samples the better to observe its statistical behaviour. The overall sample comprises 387 subjects distributed across four sub-samples: nuns, seminarians, university students and teachers. On both the statistical analysis of each sample and also on the global analysis, the scale presents two factors, denominated in the original by the terms 'identification' and 'introjection'. This bifactorial scale showed reasonable psychometric properties. Taking into account its acceptable statistical performance, we proceeded with a variance analysis by sex, where no significant differences were found, and by groups, students presenting a lower score on 'identification' and nuns a higher score on 'introjection'. The diverse results are then discussed in this paper, followed by some suggestions concerning the factorial structure of this scale and also the terminology used to denominate the two factors.

KEY-WORDS: *Religion, assessment, scale, internalization, motivation.*

Apêndice

ESCALA DE INTERIORIZAÇÃO RELIGIOSA (Ryan, Rigby e King, 1993) (traduzida e adaptada por Ferreira e Neto, 2002; readaptada por Barros, 2005)

Responda sinceramente (o questionário é anónimo) a todas as questões.

Assinale um dos números, em cada pergunta, conforme o significado seguinte:

1. *Totalmente em desacordo*
2. *Bastante em desacordo*
3. *Bastante de acordo*
4. *Totalmente de acordo*

Uma das razões por que partilho activamente a minha fé com os outros é:

- | | | | | |
|---|---|---|---|---|
| 1. porque Deus é importante para mim e eu gostaria que as outras pessoas também O conhecessem | 1 | 2 | 3 | 4 |
| (2. porque me sentiria mal comigo mesmo se o não fizesse) | 1 | 2 | 3 | 4 |
| 3. porque quero ser aceite pelos outros cristãos | 1 | 2 | 3 | 4 |

Quando me volto para Deus, faço-o, na maior parte dos casos:

- | | | | | |
|--|---|---|---|---|
| 4. porque gosto de passar tempo com Ele | 1 | 2 | 3 | 4 |
| 5. porque me sentiria culpado se o não fizesse | 1 | 2 | 3 | 4 |
| 6. porque considero que é gratificante | 1 | 2 | 3 | 4 |

A razão por que rezo sozinho é:

- | | | | | |
|--|---|---|---|---|
| 7. porque se não o faço, não serei aceite por Deus | 1 | 2 | 3 | 4 |
| 8. porque gosto de rezar | 1 | 2 | 3 | 4 |
| 9. porque considero gratificante rezar | 1 | 2 | 3 | 4 |

A razão por que frequento a igreja é:

- | | | | | |
|--|---|---|---|---|
| (10. porque se supõe que deva ir) | 1 | 2 | 3 | 4 |
| (11. porque quando vou à igreja aprendo novas coisas) | 1 | 2 | 3 | 4 |
| 12. porque os outros não me aceitariam se eu não fosse | 1 | 2 | 3 | 4 |

Os itens **1,2,4,6,8,9,11** indicam o 1º factor ("identificação").

Os itens **3,5,7,10,12** indicam o 2º factor ("introjecção").

(Os itens entre parênteses e não em negro, podem ser omitidos).

WE NEED MORE THAN SELF-REPORTS... CONTRIBUTO PARA A REFLEXÃO SOBRE AS ESTRATÉGIAS DE AVALIAÇÃO DO BEM-ESTAR

Rosa Ferreira Novo

Faculdade de Psicologia e de Ciências da Educação, Universidade de Lisboa, Portugal

Resumo

Apesar de inúmera investigação, poucos estudos têm questionado os significados e as estratégias de avaliação dos conceitos de Bem-Estar Subjectivo (BES) e Bem-Estar Psicológico (BEP). Neste trabalho discute-se a necessidade de, neste domínio, a avaliação integrar diferentes indicadores e não apenas a auto-avaliação. Esta discussão assenta numa investigação com mulheres idosas (65-75 anos), não-pacientes, e segue uma metodologia diferencial de estudo de ‘grupos experimentais’ (Reuchlin, 1964). Com recurso às EBEP (Ryff, 1989a) e ao Rorschach (Exner, 1995), foram identificados diferentes *níveis* e *tipos* de BE. A auto-avaliação mostra-se insuficiente para diferenciar tipos específicos de BE, designadamente o BEP genuíno – nível elevado de BEP na auto-avaliação e ausência de sinais psicopatológicos no Rorschach – do BEP ilusório (nível elevado de BEP na auto-avaliação e sinais de patologia e/ou sofrimento psicológico).

PALAVRAS-CHAVE: *Bem-Estar Psicológico, Bem-Estar Subjectivo, auto-avaliação, Rorschach, adultos idosos.*

Introdução

Filósofos e pensadores ocuparam-se durante séculos com o Bem-Estar (BE), tema a que a psicologia só recentemente se dedicou. Sobretudo na última década, e a par de outros temas como o optimismo, a satisfação de vida, a responsabilidade e a sabedoria, sob um domínio a que Seligman e Scizsentmihalyi (2000) designaram de Psicologia Positiva, o Bem-Estar tem sido objecto de intensa investigação. A atenção actualmente dispensada a este domínio decorre da importância científica dos temas, ao nível da investigação funda-

Morada (address): F. P. C. E. - U. L., Alameda da Universidade 1649-013 Lisboa, Portugal. E-mail: rnov@fpce.ul.pt.

Agradeço ao Prof. Doutor Mário Simões o amável convite que me dirigiu para integrar este número da Revista.

Trabalho realizado com o apoio do Centro de Psicomетria e de Psicologia da Educação da Universidade de Lisboa, Unidade I & D da Fundação para a Ciência e a Tecnologia do Ministério da Ciência, Tecnologia e Ensino Superior.

mental e da intervenção social e clínica, mas decorre certamente, também, da relevância que ele adquiriu na sociedade em geral.

Eleger como alvo de estudo dimensões que reflectem modos positivos ou de excelência do funcionamento resulta do reconhecimento de que para promover a saúde mental dos indivíduos é necessário ir além das tentativas de curar patologias instaladas ou minorar os seus efeitos. É necessário, igualmente, promover o desenvolvimento de recursos necessários à adaptação e à transformação dos ambientes e estilos de vida, à criatividade, ao envolvimento emocional e social dos indivíduos, à promoção do sentido da competência e da responsabilidade dos membros activos e produtivos de uma sociedade. Esta perspectiva de valorização do potencial de desenvolvimento individual mostra-se não só consentânea com os propósitos da psicologia, mas também se revela com grande valor ecológico. As exigências das sociedades em aceleração mudana põem continuamente à prova o bem-estar individual e, admite-se, que será o contínuo desenvolvimento humano, social e cultural que permitirá enfrentar tais exigências com maiores ganhos e menores custos individuais e sociais.

Contudo, não será propriamente este sentido de BE que tem justificado a progressiva relevância que, no domínio público, o tema veio a ter. Se no passado o interesse pelo bem-estar era veiculado pela filosofia e estava ligado à felicidade e às virtudes humanas que a condicionavam, o bem-estar veio a ser encarado, sobretudo, na sua vertente material e tornou-se um desejo de todos. Isto mesmo já denunciava Alexis Tocqueville (1805-1859) que via este interesse associado à emergência das sociedades democráticas modernas. Este político e aristocrata francês, que aos 25 anos conhece os EUA e sobre a sua gente e instituições, escreveu a obra *Democracy in America* onde considera que a procura de bem-estar centrada no plano material era a paixão dos norte-americanos. Ela levava a que os indivíduos se esforçassem por satisfazer as mais ínfimas necessidades do corpo e por alcançar as mais pequenas conveniências da vida.

Com relevância para a diferenciação dos actuais conceitos de BE adiante abordados, Tocqueville refere-se, ainda, ao facto de, no seu entender, a excessiva preocupação com o bem-estar que assenta na procura de uma gratificação imediata poder prejudicar esse mesmo bem-estar. Em última análise, é a ética subjacente à gratificação diferida das necessidades humanas que poderá conduzir à prosperidade e essa ética, segundo aquele autor, deve estar ligada à dimensão espiritual.

Pode compreender-se, à luz de Tocqueville, que a actual profusão dos temas ligados à felicidade, nas sociedades democráticas, ultrapassa de

longe o âmbito da ciência. Embora de modos diferentes, ele interessa a todos os cidadãos comuns e também aos governantes que tomam os indicadores de satisfação e felicidade como sinais para monitorizar as suas políticas.

Estas vertentes, a científica e a sócio-política, entroncam-se com as duas perspectivas conceptuais dominantes em torno do BE no âmbito da psicologia. Uma designada de Bem-Estar Subjectivo (BES) centra-se, numa perspectiva *hedónica*, na identificação do nível de felicidade e satisfação dos indivíduos e procura identificar as condições sócio-demográficas, políticas e culturais que lhes estão associadas. Outra, designada de Bem-Estar Psicológico (BEP), está enraizada no pensamento clássico desenvolvido por Aristóteles relativamente à *eudaimonia* ou felicidade. Esta confunde-se com ‘a actividade da alma dirigida pela virtude’, com a procura da perfeição e realização pessoal do *daimon* ou verdadeiro self. A perspectiva de BE baseada na *eudaimonia* insere-se na procura do *desenvolvimento do ser* e na ética subjacente à procura de gratificação diferida; de modo bem diferente, a perspectiva *hedónica* está centrada na *satisfação* do ter e associada à gratificação imediata.

Não serão inconsequentes os valores que subjazem a cada um destes tipos de Bem-Estar nem ingénua a utilização que deles pode ser feita. Não se pode escamotear a potencial influência desses valores ao nível da educação ou da organização da vida social, bem como no desenvolvimento psicológico e na saúde mental dos indivíduos. Contudo, o nosso interesse presente é, tão só, partir das diferenças conceptuais e questionar a estratégia avaliativa que ambas as perspectivas preconizam: a auto-avaliação. O estudo empírico que se apresenta permite defender a ideia de que a utilização exclusiva de escalas de auto-avaliação limita a possibilidade de alcançar a essência do conceito no âmbito da psicologia.

Conceitos de Bem-Estar

Concebido como um agregado de crenças e afectos ligados Felicidade e à Satisfação com a Vida, o construto de BES desenvolveu-se, com um progressivo dinamismo, a partir da década de 60 do séc. XX (Bradburn, 1969) e, sobretudo, nos EUA. O construto de BEP foi proposto duas décadas mais tarde, por C. Ryff (1989b), com o propósito específico de centrar este tema no âmbito do funcionamento positivo e da saúde mental e de identificar as dimensões psicológicas básicas que constituem os seus alicerces.

Embora partilhem interesses comuns, estes dois construtos têm origens distintas e conhecem orientações e percursos diferentes. O primeiro, o BES, nasce num contexto puramente empírico e tem 'navegado' sob a bandeira da felicidade e dos seus correlatos ao nível da qualidade e da satisfação de vida. Tem como critério a avaliação que as pessoas fazem das suas vidas em termos de experiências emocionais e tendo por base valores, necessidades e sentimentos pessoais. O segundo, o BEP, nasce num contexto teórico eclético que cruza a psicologia clínica e do desenvolvimento. Define como objectivo fundamental a operacionalização de dimensões do funcionamento psicológico positivo relativas ao crescimento pessoal e ao envolvimento interpessoal, dimensões que congregam características essenciais ao desenvolvimento da personalidade e à saúde mental identificadas a partir dos trabalhos de autores como C. Jung, Rogers, Maslow, Allport, Neugarten e Jahoda.

Na comparação dos dois construtos poderemos salientar que a felicidade e a saúde mental não só têm ênfases diferentes, como também assumem sentidos distintos em cada um deles. No caso do BES, a Felicidade é considerada como uma motivação fundamental da vida humana e é um critério inquestionável de BE. No caso do BEP, a felicidade é um componente e não critério. A felicidade não é um objectivo em si mesmo, mas acompanha o desenvolvimento e o funcionamento positivo num conjunto de áreas de vida (pessoal, interpessoal e social) e de domínios de funcionamento (cognitivo, afectivo-emocional e motivacional).

Na tentativa de aprofundamento do conceito de felicidade, Waterman (1993) propõe a distinção entre a felicidade como expressividade pessoal (*eudaimonia*) e como satisfação hedónica (*hedonic enjoyment*). Partindo do construto global de *self-realization*, como um domínio do funcionamento psicológico positivo, o autor conclui que a satisfação hedónica está mais ligada às actividades exercidas, aos sentimentos agradáveis no decurso da sua realização e à qualidade do desempenho atingido, ao passo que a expressividade pessoal está associada aos esforços dirigidos ao desenvolvimento pessoal, aos progressos e às expectativas de alcançar objectivos significativos face ao potencial individual. Esta proposta aplicada aos construtos em apreço permite reconhecer que o BEP abrange as duas componentes relativas à felicidade, enquanto o de BES, quer ao nível da conceptualização quer da operacionalização (ver Diener, 2000; Diener, Sapyta e Suh, 1998), se mostra associado apenas à satisfação hedónica (Compton, Smith, Cornish e Qualls, 1996; Ryff e Essex, 1991).

Medidas de Bem-Estar

Os conceitos Satisfação de Vida (*Life Satisfaction*) e com a Vida (*Satisfaction with Life*), Optimismo (*Optimism*), Felicidade (*Happiness*), Moral (*Morale*), Afecto Positivo (*Positive Affect*), Afecto Negativo (*Negative Affect*), Equilíbrio Afectivo (*Affect Balance*), Depressão (*Depression*), Bem-Estar Subjectivo (*Subjective Well-Being; Welfare Subjective*) e Qualidade de Vida (*Life Quality e Perceived Life Quality*), entre outros, são utilizados diversamente como indicadores de BES global ou de um das suas componentes específicas e, por vezes ainda, como indicadores de BEP. A esta proliferação de termos, apenas potencialmente equivalentes, acresce a existência de diferentes medidas para cada um deles e a escassez de estudos de validade das medidas.

Na última década, a Satisfação com a Vida e o de Felicidade afirmaram-se como indicadores dominantes de BES. A diversidade de instrumentos e de utilização das medidas finais por parte de diferentes investigadores é, contudo, surpreendente. É exemplo a dimensão *felicidade*, e que, mesmo quando é 'medida' através do afecto positivo e do afecto negativo, o modo como os resultados destas duas componentes são ponderados, combinados ou diferenciados, varia de estudo para estudo. Numa avaliação das medidas de BES, Andrews e Robinson (1991) concluem que há inúmeras 'escalas' com apenas um item e, entre as multi-itens, "não há uma escala, ou mesmo um conjunto pequeno de escalas, que seja especialmente usada ou claramente melhor que as outras" (p. 70). A *Delighted-Terrible Scale*, de Andrews e Withey (1976, cit. in Andrews e Robinson, 1991), é uma escala composta por um item apenas ("Como é que se sente acerca da sua vida em geral") ao qual é solicitada uma resposta num contínuo de sete pontos ordenados de 'satisfeito/contente' a 'insatisfeito/descontente'. Outras medidas de avaliação do BES integram vários itens. Assim, a escala de Afecto Positivo e Negativo (*Positive and Negative Affect Scale* – PANAS, de Watson, Clark e Tellegen, 1988), por exemplo, avalia as duas dimensões do afecto, através de 10 itens para cada uma das dimensões, numa escala de tipo Likert em cinco pontos; a escala de Satisfação com a Vida (*Satisfaction With Life Scale*; Diener, Emmons, Larsen e Griffin, 1985) avalia o acordo ou desacordo dos indivíduos, numa escala de sete pontos, com um conjunto de cinco itens apenas; a escala de felicidade subjectiva (*Subjective Happiness Scale*) de Lyubomirsky e Lepper (1999, cit. in Lyubomirsky, 2001) integra apenas quatro itens.

Apesar de uma evolução no sentido de indicadores mais amplos e representativos dos construtos em causa, as medidas mais restritas e pobres do ponto de vista psicométrico continuam a ser usadas. Compreensível o seu uso

em inquéritos de carácter sociológico e/ou transcultural que envolvem grandes amostras (e.g., *General Social Surveys*, conduzidos nos EUA; *European Social Surveys*, na Europa), mas menos útil em estudos de outra natureza.

Por outro lado, a diversidade das medidas tem consequências no plano teórico, sobretudo porque a construção teórica no âmbito do BES tem sido feita a partir da integração de dados empíricos (*bottom-up empirical way*) (Diener et al., 1998). Esta opção, todavia, deveria ser acompanhada de maior exigência na selecção das medidas e dos critérios de operacionalização do construto.

Referindo-se às medidas de BES, Schwarz e Strack (1999) chamam a atenção para o facto da avaliação do nível de BE ser função: do tipo de escalas utilizadas, da ordem em que os itens são apresentados, do foco temporal visado em cada escala, dos factores situacionais e contextuais da avaliação, do estado específico de humor no momento da avaliação e de factores de desejabilidade social. Diener e colegas (Diener, Suh, Lucas, e Smith, 1999) consideram que os factores situacionais são de menor relevo em comparação com a influência a longo prazo das medidas de BES. Por outro lado, Diener (2000) contorna a questão da desejabilidade social, salientando o facto de as questões visarem a experiência subjectiva avaliada por processos de comparação social: os indicadores avaliam se as pessoas se sentem mais ou menos felizes que os seus pares. Contudo, independentemente de se visar a experiência subjectiva, o problema da avaliação neste domínio é relativo ao acesso aos dados da própria experiência subjectiva e aos critérios e significados que cada um atribui aos conceitos em causa.

Como síntese, os aspectos críticos assinalados não põem em causa o rigor metodológico de cada um dos trabalhos de *per se*, mas, antes, assinalam as dificuldades de construção teórica neste domínio, pelo menor poder cumulativo da informação proveniente de investigação conduzida com o recurso a metodologias e critérios bastante diversos.

Relativamente às medidas de BEP elas foram desenvolvidas a par da construção teórica (Ryff, 1985, 1989a,b). As teorias em que se alicerça o conceito apresentavam já definições de personalidade 'ideal', 'madura' ou 'realizada' que, embora amplas e difusas, sinalizam as características essenciais de funcionamento positivo. O mérito de Ryff foi o de identificar os pontos de convergência das diversas formulações e de as transportar para o plano empírico. O modelo síntese de BE que veio a propor indica seis dimensões nucleares do desenvolvimento normal, num sentido positivo e de bem-estar, a saber: Aceitação de Si, Relações Positivas com os Outros, Domínio do Meio, Crescimento Pessoal, Objectivos na Vida e Autonomia.

Para cada uma destas dimensões foi proposta uma escala específica de auto-avaliação, de tipo Likert, em que os itens são afirmações de carácter descritivo com seis categorias de resposta ordenada do Discordo Completamente ao Concorde Completamente. A versão final do instrumento integra o conjunto das seis escalas designadas daqui em diante como Escalas de Bem-Estar Psicológico (EBEP). Existem actualmente três versões destas escalas que diferem apenas no número de itens: a primeira versão proposta integra vinte itens por dimensão, sendo o instrumento global de 120 itens; a segunda e terceira versões são constituídas por catorze e três itens por dimensão, num total de 84 e de 18 itens, respectivamente. Estas últimas duas versões são, actualmente, as mais utilizadas.

Ao conceber o BE de uma forma abrangente, como o resultado de diversos processos cognitivos, afectivos e emocionais, o modelo de BEP constitui uma via integradora da definição e avaliação do construto. O modelo proposto por Ryff permite descrever amplas dimensões da vivência psicológica subjectiva: a relação da pessoa consigo própria e com a sua vida no presente e no passado, a capacidade para definir e orientar a vida em função de objectivos significativos para si própria, a natureza e qualidade da relação com o meio intersubjectivo e social. Do ponto de vista da fundamentação teórica das escalas e do processo de validação das medidas, este modelo do BEP oferece também indicadores específicos mais sólidos, robustos e consistentes do que os de BES.

Limitações na avaliação do Bem-Estar

Os modelos de avaliação do BES e BEP, como vimos, distinguem-se pela amplitude dos domínios avaliados e aproximam-se quanto à modalidade de avaliação que preconizam. Eles têm em comum a possibilidade de caracterizar a percepção pessoal de felicidade e de satisfação com as experiências de vida, tendo por referência os critérios do próprio indivíduo e, como limitação, as suas capacidades de auto-análise e o acesso a informação potencialmente relevante. O Bem-Estar é, em qualquer dos modelos, a expressão pessoal e subjectiva, a avaliação directa e introspectiva com base nos valores e padrões pessoais.

Mas será que algum dos modelos, seguindo a estratégia de auto-avaliação, mesmo que numa vasta área de conteúdos (relativos a crenças, sentimentos ou comportamentos) garante um funcionamento psicológico positivo e constitui um indicador de genuína saúde mental? Somos de parecer

que não: o modelo do BEP porque envolve limitações metodológicas e o do BES porque, além das metodológicas, envolve também limitações conceptuais.

Um funcionamento psicológico harmonioso e positivo constitui, teoricamente, uma condição prévia de BE, designadamente de BEP. Será a auto-avaliação capaz de captar a qualidade do funcionamento psicológico? A normalidade de funcionamento será avaliável apenas no plano da vivência subjectiva?

O BE avaliado a partir dos conteúdos da experiência subjectiva dá a conhecer uma das faces do funcionamento. A outra face, a qualidade do funcionamento que é correlativa da saúde mental, só pode ser avaliada a partir dos processos envolvidos na experiência subjectiva e não-subjectiva, do pensamento, do sentimento e do comportamento. É a qualidade dos *processos* e não a natureza dos *conteúdos* das vivências que permitirá identificar a normalidade ou a saúde mental. As experiências subjectivas podem (e devem) constituir objecto de estudo psicológico, mas não podem constituir-se como critérios de saúde mental.

Se a conceptualização do BEP em termos fenomenológicos exige o acesso ao mundo subjectivo, pelo olhar do próprio, a sua conceptualização em termos de saúde mental exige o acesso às manifestações desse BE, pelo olhar externo. Diz-nos Miranda Santos (2002) que a observação do ser humano pelo humano nos confronta com os limites da ciência objectiva para compreender o que é radicalmente subjectivo. O desafio, neste contexto, é o da exigência de uma perspectiva que conjugue a avaliação interna e subjectiva e a avaliação externa, funcional e objectiva. A complementaridade destas duas perspectivas de avaliação constitui uma via de análise científica objectiva das experiências subjectivas.

É esta questão que é visada no trabalho empírico que adiante se apresenta. Defende-se a necessidade de, neste domínio de estudo, conjugar as vantagens da auto-avaliação com as técnicas projectivas e, assim, ultrapassar as limitações específicas a cada uma destas metodologias de *per se*. No que se refere ao BEP, não devemos limitar-nos aos conteúdos que emanam da reflexão e auto-análise por parte de cada um sobre si próprio, mas integrar os processos psicológicos que sustentam o comportamento expressivo nas suas múltiplas vertentes. As formas próprias de cada um processar a informação, atribuir sentido e construir significados constitui uma via de acesso ao mundo privado e é, potencialmente, uma forma privilegiada de avaliação objectiva do mundo interno. Criar condições para que este funcionamento se revele, isto é, potenciar a exteriorização dos processos internos, e identificar parâmetros

objectivos de análise das manifestações de tais processos – são duas ambições das técnicas projectivas – constituem uma forma de aceder ao mundo privado, aos conteúdos e aos processos que estão na base do comportamento adaptativo.

Estudo Empírico

Objectivos

Se se assume que o BEP está, por um lado, intimamente relacionado com o desenvolvimento da personalidade (Ryff, 1989a) e, por outro, se inscreve num processo de funcionamento psicológico positivo (Ryff, 1989a,b), impõe-se a necessidade de alargar as estratégias de avaliação. Neste sentido, o estudo empírico que se apresenta - e que constitui uma parcela de uma investigação empreendida com o propósito mais vasto de clarificar a natureza e a dinâmica do BEP no âmbito da personalidade (Novo, 2000/2003) - contribui para nos confrontar com as insuficiências da metodologia de auto-avaliação neste domínio. Pretende-se defender a ideia de que, para além das medidas indirectas e de auto-avaliação comumente utilizadas, é necessário integrar medidas comportamentais e de avaliação externa, capazes de captar a qualidade do funcionamento psicológico.

Metodologia

Foi definida a hipótese de que *níveis* elevados de BEP, expressos na auto-avaliação, não indicam necessariamente um BEP genuíno. Por BEP genuíno entende-se a auto-avaliação de bem-estar num contexto de funcionamento psicológico positivo e de saúde mental. O teste da hipótese seguiu uma metodologia diferencial de comparação entre grupos 'experimentais' representativos de 'grupos teóricos', isto é, organizados a partir de um critério abstracto, fundamentado teoricamente e operacionalizado empiricamente (Reuchlin, 1964).

Fazendo jus aos fundamentos do construto, é considerado como critério teórico de *BEP genuíno* a presença de cognições e emoções de bem-estar, mas também a ausência de sinais de patologia mental e/ ou um sofrimento psicológico. No plano empírico, este critério é operacionalizado a partir da conjugação dos dados das EBEP (nível de BEP: reduzido ou elevado) e dos indica-

dores do Rorschach de risco de patologia e/ou de vulnerabilidade psicológica (ausência ou presença).

Tal critério permitirá prever a existência de vários *tipos* de BEP. Particularmente atentos à auto-avaliação positiva, considera-se com BEP genuíno as participantes com níveis elevados na auto-avaliação e sem sinais de patologia ou de risco psicológico. As participantes com resultados elevados na auto-avaliação e com sinais de patologia ou vulnerabilidade serão consideradas com BEP ilusório. Os restantes casos, com níveis inferiores à média na auto-avaliação de BEP, serão considerados com BEP reduzido.

Amostra

Foi estudada uma amostra de 69 mulheres, com idades entre os 65 e os 75 anos, inseridas satisfatoriamente num meio urbano e em condições de normalidade física e psíquica. Por razões teóricas e relacionadas com o desafio que o envelhecimento representa em termos da construção do BE, a idade constituiu um critério primário na selecção da amostra; o sexo foi um critério secundário, pelo que a opção pelo feminino decorreu do interesse e disponibilidade de participantes.

As participantes foram recrutadas em instituições culturais, comunitárias e de saúde da área da Grande Lisboa e foram seleccionadas entre 120 participantes numa primeira fase do estudo. Após exclusão de alguns casos (doenças crónicas, suspeitas de patologia neurológica, história psiquiátrica e acontecimentos de vida críticos no último ano) a selecção foi feita em função dos critérios de constituição dos grupos 'experimentais', designadamente do resultado obtido nas EBEP.

No Quadro 1 é apresentada a síntese das características sócio-demográficas da amostra. A média de idade das participantes é de 68 anos e o nível de escolaridade tem a mediana nos nove anos, superior ao da população portuguesa para coortes correspondentes. O casamento e a viuvez são os estados civis mais representados e as participantes são aposentadas, com excepção de duas que se mantêm, por opção, na vida profissional activa.

Quadro 1 - Características sócio-demográficas da amostra

	FREQUÊNCIA	%
IDADE (anos)		
65	23	33%
66 – 70	25	36%
71 – 75	21	30%
NÍVEL DE ESCOLARIDADE (anos)		
4	20	29%
5 – 10	22	32%
11 – 12	17	25%
> 12	10	14%
ESTADO CIVIL		
Solteiro	9	13%
Casado	24	35%
Divórcio/Separação	13	19%
Viúvo	23	33%

Nota. N = 69; Sexo: Feminino

Instrumentos

Orientada pela necessidade de focar as dimensões em estudo e de representar estratégias de avaliação distintas, a selecção de instrumentos recaiu nas EBEP e no Rorschach.

Escala de Bem-Estar Psicológico (EBEP). Foi utilizada a versão portuguesa com 14 itens por cada uma das seis dimensões avaliadas, formando, o conjunto de 84 itens, um instrumento único (ver Novo, Duarte Silva e Peralta, 1997; Novo 2000/2003). Na versão original, os índices de consistência interna (*alpha* de Cronbach) variam entre .86 e .93, e a estabilidade temporal (coef. de Pearson) entre .81 e .88 (Ryff e Essex, 1991). Na versão portuguesa, os coeficientes *alpha* variam entre .74 e .86, e a estabilidade temporal (*tau* de Kendall) para as seis medidas situa-se entre .41 e .83 (Novo et al., 1997). Muito embora não tenha sido considerada na versão original, uma medida global para o conjunto das seis dimensões foi proposta e fundamentada, revelando-se psicometricamente mais robusta do que os indicadores por dimensão (estabilidade temporal .81; consistência interna .93) e útil na diferenciação de níveis de BEP (Novo 2000/2003).

Sistema Integrativo de Rorschach (S.I.R). Foi seguida a metodologia proposta no Workbook for the Comprehensive System (Exner, 1995). O Rorschach, instrumento com características específicas que o tornam um dos mais utilizados na avaliação clínica e na investigação de dimensões estruturais e di-

nâmicas da personalidade, integra escalas e índices validados empiricamente e capazes de se constituir como indicadores de vulnerabilidade e patologia. É ainda um instrumento com baixa susceptibilidade à falsificação consciente e à distorção defensiva (Viglione e Rivera, 2003) e revela-se com características metrológicas de precisão e validade adequadas (Weiner, 1998; ver estudos de validade em Exner, 1991, 1993; Exner e Weiner, 1995; Weiner e Exner, 1991). No presente estudo foi testada a precisão das codificações, através de uma segunda cotação 'cega' de 30% dos protocolos (selecção aleatória), tendo o grau de acordo intercotadores sido superior a 90%.

Procedimento

Numa primeira fase, foi determinado o resultado global nas EBEP e identificadas as participantes com nível reduzido e elevado de BEP, isto é, com resultados abaixo e acima da média, respectivamente, de acordo com os valores de referência para a população portuguesa (ver Novo, 2000/2003). Posteriormente foram analisados os Sumários Estruturais do Rorschach e calculados os índices indicadores de potencial patologia e/ou vulnerabilidade.

Os indicadores considerados foram aqueles que, estando validados empiricamente, sinalizam alterações do funcionamento, dificuldades de adaptação ou sofrimento psicológico, designadamente: *SCZI* (índice de esquizofrenia); *WSum6* (índice de rastreio de alterações da lógica e coerência do pensamento); *DEPI* (índice de depressão e de outras desordens afectivas); *S-Con* (constelação de suicídio); *Nota Adj D* (índice de desordens de ansiedade e de alterações crónicas de controlo e tolerância ao stress); *X+%*, *Xu%* e *X-%* (variáveis de rastreio de alterações da percepção). Foi definido como critério de vulnerabilidade psicológica a existência de pelo menos um dos seguintes indicadores:

- *SCZI* > 4 ou *WSum 6* > 11
- *S-Con* > 7
- *DEPI* > 4
- *Nota AjdD* < -1
- *X+%* < .61 e *X-%* ≥ *Xu%*

Resultados

A análise de resultados é apresentada em duas fases: a primeira, centrada exclusivamente nos resultados da auto-avaliação, visa a identificação de diferentes níveis de BEP; a segunda, é orientada para a procura de convergência ou divergência entre os dados da auto-avaliação e do Rorschach, e vi-

sa a constituição de grupos representativos de diferentes *tipos* de BEP e a análise dos dados relevantes.

Níveis de BEP

Tendo por base o Resultado Total de BEP do estudo de adaptação à população portuguesa das escalas (ver Novo, 2000/2003), designadamente a média (e o erro-padrão da média) do grupo de idade correspondente à presente amostra, procedeu-se à identificação dos pontos de corte que permitissem a diferenciação de dois níveis de BEP: abaixo da média e acima dela. Foram assim seleccionadas as 69 participantes que vieram a integrar o estudo e constituídos dois grupos com diferentes 'níveis' de Bem-Estar Psicológico (entre o valor máximo do grupo abaixo da média e o valor mínimo do grupo acima da média, distam cerca de dez erros-padrão da média). O grupo com resultados abaixo da média é constituído pelas 16 participantes com Resultado Total de BEP (23% da amostra) com valores entre 268 e 335 ($M = 308.25$; $SD = 21.49$); o segundo grupo com 53 participantes (77% da amostra) com valores entre 365 e 462 ($M = 396.96$; $SD = 27.31$).

Tipos de BEP

Realizada a análise e identificação dos sinais de vulnerabilidade psicológica, foram sinalizados 39 casos (a vulnerabilidade psicológica identificada não remete para qualquer diagnóstico, dado que não se procedeu a uma avaliação psicológica global; ela apenas sinaliza alterações específicas a cada caso e que, portanto, não tipificam o grupo em si). Dos casos sinalizados, 14 são participantes já identificadas com nível reduzido de BEP e outros 25 são relativos a participantes com nível elevado de BEP.

Apesar do cruzamento dos dois indicadores (EBEP e Rorschach) tornar possível a constituição de quatro grupos com tipos de BEP distintos, apenas três nos pareceram relevantes. A quase totalidade das participantes com reduzido BEP (14 dos 16 casos) revelou sinais de vulnerabilidade psicológica; apenas duas das participantes não apresentarem tais sinais. Assim, as participantes com nível de BEP abaixo da média constituem um só grupo ($n = 16$) que será identificado como Grupo 1 (BEP-R) – Grupo de Bem-Estar Psicológico Reduzido.

No que se refere às 53 participantes com *nível* elevado de BEP, há lugar à constituição de dois grupos. Um que agrega as 25 participantes com auto-avaliação positiva de BE e com sinais de risco ou vulnerabilidade psicológica, grupo que será identificado como Grupo 2 (BEP-VP) – Grupo de Bem-Estar e

Vulnerabilidade Psicológica. O outro grupo agrega as restantes 28 participantes sem de sinais de vulnerabilidade e que satisfazem o critério de BEP genuíno; este grupo será identificado como Grupo 3 (BEP-G) – Grupo de Bem-Estar Psicológico Global.

Numa leitura dos resultados das EBEP (ver Quadro 2) salientam-se: diferenças inter-grupos em quatro dimensões de BEP; diferenças que ocorrem sempre entre o Grupo 1 e os Grupos 2 e 3 que, por sua vez, não se diferenciavam entre si (ver Fig. 1) – a diferenciação entre dois últimos grupos só é possível a partir dos dados do Rorschach.

Quadro 2 - Médias e Desvios-Padrão dos resultados das escalas de Bem-Estar Psicológico

MEDIDAS DE BEM-ESTAR PSICOLÓGICO	AMOSTRA TOTAL (N = 69) M (SD)	GRUPO 1 (n = 16)	GRUPO 2 (n = 25)	GRUPO 3 (n = 28)	H ^a (2, N = 69)	Comparações Múltiplas
Autonomia	63.41 (9.76)	56.87 (13.26)	65.24 (8.75)	65.50 (6.43)	5.68 (ns)	—
Domínio do Meio	61.88 (8.68)	51.94 (5.87)	64.68 (6.36)	65.07 (7.63)	26.35***	G1 < G2*** G1 < G3***
Crescimento Pessoal	61.93 (12.02)	55.37 (12.61)	63.64 (10.90)	64.14 (11.69)	5.01 (ns)	—
Relações Positivas	65.99 (12.71)	51.06 (11.23)	68.80 (11.40)	72.00 (6.69)	23.47***	G1 < G2*** G1 < G3***
Objectivos Vida	60.83 (11.60)	46.50 (9.23)	63.84 (8.68)	66.32 (7.97)	29.32***	G1 < G2*** G1 < G3***
Aceitação de Si	62.36 (12.65)	45.50 (11.32)	66.92 (8.99)	67.36 (8.10)	28.69***	G1 < G2*** G1 < G3***
RESULTADO TOTAL	376.39 (45.77)	308.25 (21.49)	393.12 (27.67)	400.39 (27.04)	36.91***	G1 < G2*** G1 < G3***

Nota. Grupo 1 – BEP-R (Bem-Estar Psicológico Reduzido); Grupo 2 – BEP-VP (Bem-Estar e Vulnerabilidade Psicológica); Grupo 3 – BEP-G (Bem-Estar Psicológico Global) a Teste de Kruskal-Wallis; *** $p < .001$

A Autonomia – expressa pela determinação para pensar e agir com base em padrões pessoais e pela capacidade para resistir às pressões sociais – e o Crescimento Pessoal – expresso pela abertura a novas experiências e a percepção de desenvolvimento – são as duas áreas de maior bem-estar entre as participantes do Grupo 1 e sinalizam características presentes com intensidade ou frequência similar às restantes participantes; é nas restantes dimensões (Domínio do Meio, Relações Positivas com os Outros, Objectivos na Vida e Aceitação de Si) que o reduzido BEP se revela significativamente inferior às participantes dos outros dois grupos.

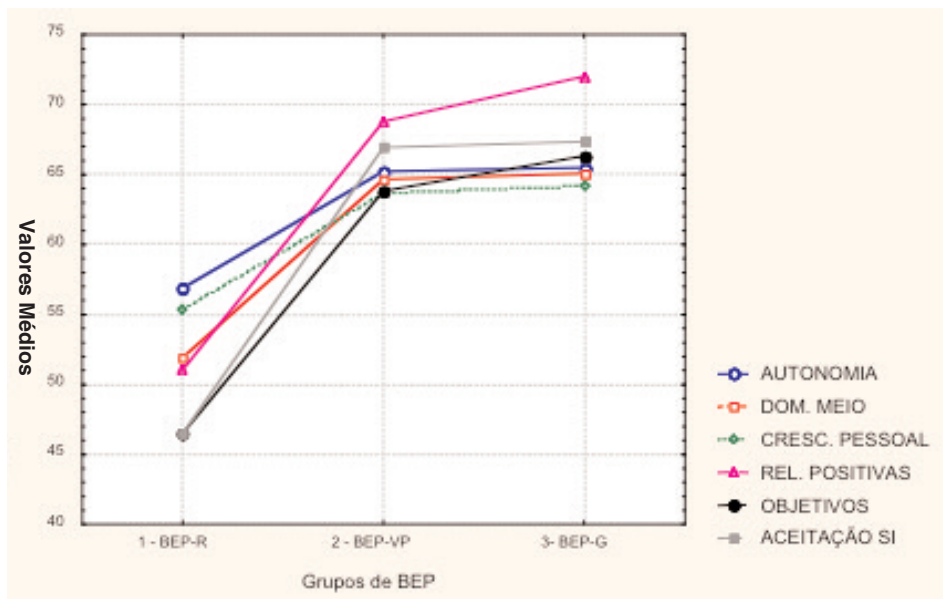


Figura 1 - Valores Médios das Escalas de Bem-Estar Psicológico

As participantes do Grupo 2 têm em comum o facto de, face aos aspectos da experiência pessoal e subjectiva, da personalidade e do comportamento, se auto-avaliarem com elevado Bem-Estar, não obstante apresentarem sinais de alteração ao normal funcionamento e harmonia psicológica. As participantes do Grupo 3, por seu lado, têm em comum uma auto-avaliação positiva, sentimentos de realização e adequação, e não apresentam indícios de vulnerabilidade ou de risco psicopatológico.

Estes resultados indicam que existem *tipos* distintos de BEP que a auto-avaliação não permite revelar. Do conjunto de 53 participantes com níveis elevados de BE, 47% apresenta sinais significativos de vulnerabilidade psicológica que comprometem a harmonia de funcionamento que o BEP pretende sinalizar. Conclui-se, assim, que uma auto-avaliação de nível elevado nem sempre traduz um funcionamento positivo e genuíno BEP.

Conclusões

O modelo do BEP proposto por C. Ryff mostra-se, do ponto de vista teórico e metodológico, mais adequado a uma exploração do tema no âmbito da psicologia do que o modelo de BES. Ele abre perspectivas diversas, no âmbito da personalidade e da saúde mental, ao nível da investigação e da interven-

ção. Ele tem também o mérito de servir de referência ao questionamento da ética subjacente às decisões políticas, nos diversos domínios sociais e educativos, tomadas em nome do desenvolvimento e do bem-estar dos cidadãos.

Contudo, a metodologia de avaliação que tem sido utilizada, porque limitada à auto-avaliação da vivência subjectiva de BE, mostra-se insuficiente. No estudo apresentado, cerca de metade dos casos de auto-avaliação positiva de BEP vêm a revelar-se, na avaliação complementar, em situação de vulnerabilidade e/ou de sofrimento psicológico. As escalas de auto-avaliação mostram-se, assim, incapazes de diferenciar entre este *tipo* de BEP ilusório (em que as dificuldades de adaptação e/ou o sofrimento psicológico não é reconhecido) e um BEP genuíno (alicerçado na saúde mental e na harmonia do funcionamento psicológico).

O enquadramento teórico e epistemológico do construto, designadamente a ênfase na normalidade do desenvolvimento e na positividade do funcionamento, reclama estratégias de avaliação consentâneas com aquela ambição. Os dados obtidos são naturalmente limitados, por um lado, a uma dada coorte e amostra muito específicas, por outro, ao uso de uma metodologia de avaliação complexa que, naturalmente, não pode ser usada por rotina. Contudo, os resultados sugerem a necessidade de considerar a convergência/divergência entre medidas de auto-avaliação e outras de avaliação externa, objectiva, funcional ou comportamental. Esta integração amplia as possibilidades de investigação, ao mesmo tempo que permite captar a essência do conceito no âmbito da psicologia e torná-lo um potencial indicador de 'saúde psicológica'.

Referências

- Andrews, F. M. e Robinson, J. P. (1991). Measures of subjective well-being. In J. P. Robinson, P. R. Shaver e L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (Vol. 1, pp. 61-114). San Diego, CA: Academic Press.
- Bradburn, N. M. (1969). *The structure of psychological well-being*. Chicago: Aldine.
- Compton, W., Smith, M., Cornish, K. e Qualls, D. L. (1996). Factor structure of mental health measures. *Journal of Personality and Social Psychology*, 71, 406-413.
- Diener, E. (2000). Subjective well-being: The science of happiness and a proposal for a national index. *American Psychologist*, 55, 34-43.
- Diener, E., Sapyta, J. e Suh, E. M. (1998). Subjective well-being is essential to well-being. *Psychological Inquiry*, 9, 33-37.
- Diener, E., Suh E., M., Lucas, R. E. e Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin*, 125, 276-302.
- Exner, J. E. (1991). *The Rorschach: A Comprehensive System. Volume 2: Interpretation* (2ª ed.). New York: John Wiley & Sons.
- Exner, J. E. (1993). *The Rorschach: A Comprehensive System. Volume 1: Basic foundations* (3ª ed.). New York: John Wiley & Sons.
- Exner, J. E. (com Colligan, S. C., Hillman, L., Ritzler, B., Sciara, A., e Viglione, D.) (1995). *A Rorschach workbook for the Comprehensive System* (4ª ed.). Asheville, North Carolina: Rorschach Workshops.
- Exner, J. e Weiner, I. (1995). *The Rorschach: A Comprehensive System. Volume 2: Interpretation* (3ª ed.). New York: John Wiley & Sons.
- Lyubomirsky, S. (2001). Why are some people happier than others? The role of cognitive and motivational processes in Well-Being. *American Psychologist*, 56, 239-249.
- Novo, R. F. (2003). *Para além da Eudaimonia: O Bem-Estar Psicológico em mulheres na idade adulta avançada*. Coimbra: Fundação Calouste Gulbenkian e Fundação para a Ciência e Tecnologia. (Trabalho original de 2000).
- Novo, R. F., Duarte-Silva, M. E., e Peralta, E. (1997). O bem-estar psicológico em adultos: Estudo das características psicométricas da versão portuguesa das escalas de C. Ryff. In M. Gonçalves, I. Ribeiro, S. Araújo, C. Machado, L. Almeida, M. Simões (Orgs.), *Avaliação psicológica: Formas e contextos* (Vol. V, pp. 313-324). Braga: Associação dos Psicólogos Portugueses.
- Reuchlin, M. (1964). Domaine et méthodes de la psychologie differentielle. *Enfance*, 17, 39-50.
- Ryff, C. (1989a). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, 57, 1069-1081.

- Ryff, C. (1989b). Beyond Ponce de Leon and life satisfaction: New directions in quest for successful aging. *International Journal of Behavioral Development*, 12, 35-55.
- Ryff, C. e Essex, M. (1991). Psychological well-being in adulthood and old age: Descriptive markers and explanatory processes. In K. Schaie e M. Lawton (Eds.), *Annual Review of Gerontology and Geriatrics* (Vol. 11, pp. 144-171). New York: Springer.
- Santos, A. M. (2002). Apostar na ciênciapsicologia. *Psychologica*, 30, 487-510.
- Schwarz, N. e Strack, F. (1999). Reports of subjective well-being: Judgmental processes and their methodological implications. In D. Kahneman, E. Diener, e N. Schwartz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 61-84). New York: Russell Sage Foundation.
- Seligman, M. P. e Csikszentmihalyi, M. (2000). Positive Psychology: An introduction. *American Psychologist*, 55, 5-14.
- Viglione, D. J. e Rivera, B. (2003). Assessing personality and psychopathology with projective methods. In I. Weiner, J. Graham e J. Naglieri (Eds.), *Handbook of psychology. Vol. 10: Assessing psychology* (pp. 531-552). New York: J. Wiley & Sons.
- Waterman, A. (1993). Two conceptions of happiness: Contrasts of personal expressiveness (eudaimonia) and hedonic enjoyment. *Journal of Personality and Social Psychology*, 64, 678-691.
- Watson, D. Clark, L. e Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54, 1063-1070.
- Weiner, I. B. (1998). *Principles of Rorschach interpretation*. Mahwah, NJ: Lawrence Erlbaum.
- Weiner, I. B. e Exner, J. (1991). Rorschach changes in long-term and short-term psychotherapy. *Journal of Personality Assessment*, 56, 453-465.

WE NEED MORE THAN SELF-REPORTS... **CONTRIBUTION FOR THE REFLECTION ABOUT WELL-BEING ASSESSMENT STRATEGIES**

Rosa Ferreira Novo

Faculdade de Psicologia e de Ciências da Educação, Universidade de Lisboa, Portugal

Abstract: Although the topic of well-being has generated considerable research, few studies have explored the meanings and the assessment strategies of the SWB and PWB concepts. In this work we discuss the necessity of the psychological assessment encompasses other than self-reports data. An empirical study with 69 older women (65-75 years), non-patients, was delineated with a differential methodology of 'experimental groups' (Reuchlin, 1964). The PWBS (Ryff, 1989a) and the Rorschach (Exner, 1995) was used to identify different levels and patterns of PWB. The results revealed that self-reports are insufficient to differentiate the genuine well-being – self-reported high levels of PWP and without signs of pathology – from the illusion of well-being – high levels of PWP tied to signs of the psychological vulnerability and/or pain).

KEY-WORDS: *Subjective Well-Being, Psychological Well-Being, self-report, Rorschach, older adults.*

PODE UM GRUPO PEDIR PERDÃO A OUTRO GRUPO? UMA PERSPECTIVA TIMORENSE

Félix Neto

Maria do Conceição Pinto

Universidade do Porto, Portugal

Etienne Mullet

Ecole Pratique des Hautes Etudes, France

Resumo

Podem as nações “enveredar por processos que têm como resultado arrependimento e perdão colectivo?” (Shriver, 1995, p. 71). O presente estudo examina esta questão mediante a análise da perspectiva de pessoas de Timor-Leste que estiveram recentemente envolvidas numa guerra de ocupação e em que a maior parte delas sofreram pessoalmente em resultado de muitos conflitos. Participaram neste estudo 226 pessoas (111 mulheres e 115 homens) com idades oscilando entre 19 e 70 anos. O material consistiu num questionário com 77 proposições a respeito de possíveis atitudes acerca de pedir perdão. Uma grande maioria dos participantes eram favoráveis à ideia de que um grupo de pessoas pode pedir perdão a um outro grupo de pessoas. Menos de 9% dos participantes consideraram que não era possível a um grupo de pessoas pedir perdão a outro grupo de pessoas. Os resultados fundamentais eram robustos no que concerne a idade, o género e outras diferenças individuais. Os resultados estão em consonância com os de estudos anteriores que examinaram o perdão interpessoal na perspectiva de dar perdão (Neto, Pinto, e Mullet, 2004).

PALAVRAS-CHAVE: *Perdão intergrupar, Timor, reconciliação.*

Introdução

Podem as nações “enveredar por processos que têm como resultado arrependimento e perdão colectivo?” (Shriver, 1995, p. 71). O presente estudo examina esta questão mediante a análise da perspectiva de pessoas de Timor-Leste que estiveram recentemente envolvidas numa guerra de ocupação e em que a maior parte delas sofreram pessoalmente em resultado de muitos conflitos.

Morada (address): Faculdade de Psicologia, Universidade do Porto, Rua Manuel Pereira da Silva, 4200-329 Porto.
E-mail: fneto@fpce.up.pt

Agradece-se o apoio do programa POCTI/PSI/46245/2002 da Fundação para a Ciência e a Tecnologia, do Laboratório Cognition et Décision da École Pratique des Hautes Études e da UMR Travail et Cognition da Universidade de Toulouse Le Mirail.

O perdão pode não parecer relevante para a ética política, pois fora durante muito tempo concebido por filósofos morais (Smedes, 1996) e subsequentemente por psicólogos sociais (para uma revisão cf. Worthington, 2005) como um processo que só pode envolver pessoalmente um ofensor e uma pessoa ofendida. A primeira tentativa para estudar o perdão em contextos socio-políticos foi levada a cabo somente ao nível do perdão interpessoal (Azar, Mullet & Vinsonneau, 1999; Azar and Mullet 2001, 2002a, 2002b).

Esta concepção do perdão não toma em consideração que muitas, senão a maior parte das ofensas na vida social são colectivas (Minow, 1998); que na guerra, em particular, as ofensas são cometidas não somente contra pessoas, mas “contra a própria sociedade” (Thomas e Garrod, 2002); que as responsabilidades são muitas vezes partilhadas por muitas pessoas (Shriver, 1995); que a própria justiça para elas é muitas vezes intratável (Digeser, 2001); que a sua confissão deve, para ser completa, ser uma empreitada colectiva; e que a sua cura só pode ser levada a cabo ao nível da comunidade. Todavia estes factores parecem ser compreendidos pelas pessoas comuns (Mullet, Girard e Bakhshi, 2004).

Shriver no seu livro de 1995 com o título *“An Ethics for Enemies: Forgiveness in Politics”* analisou numerosos exemplos de acções de líderes políticos que encaixam nas categorias de arrependimento e de perdão. Como referiu o autor “Se os líderes acusam um inimigo de crime, a confissão para crimes das suas próprias pessoas, ou oferecer esperanças para uma futura reconciliação, fazem tudo isto em nome de um colectivo em relação a outro. A negação deste papel representativo e simbólico aos políticos constitui um empobrecimento do seu serviço num tratamento de uma sociedade com as suas ofensas passadas e a sua actual responsabilidade correctiva para o futuro” (Shriver, 1995, p. 113). Ele citou figuras políticas contemporâneas, tais como Martin Luther King, o chanceler alemão Willy Brandt e o presidente Richard von Weizsäcker, os primeiros ministros Kiichi Miyazawa e Morihiro Hosokawa, e figuras políticas do passado, tais como os líderes da Convenção Charleston dos Povos de Cor de 1985. Se Shriver escrevesse hoje este livro poderia certamente acrescentar a esta lista o Presidente da África do Sul Nelson Mandela (ver também Henderson, 1996, para numerosos outros exemplos de perdão em política).

Objectivos do presente estudo

O objectivo do presente estudo foi o de examinar em pormenor as perspectivas e atitudes de pessoas comuns em relação ao significado do perdão

intergrupar e o modo como o perdão intergrupar, se julgado imbuído de sentido, podia ocorrer. Só foi considerada uma das facetas do perdão intergrupar: pedir perdão. Dar perdão, se foi ou não pedido, já foi examinado em estudo anterior (Neto, Pinto e Mullet, 2004). O leitor interessado em conhecer o contexto histórico do estudo presente pode consultar esse artigo para ter mais informações.

Nesse estudo, vários aspectos concretos do processo de dar perdão, foram examinados, entre eles: O perdão intergrupar tem sentido *in abstracto*? O perdão intergrupar faz sentido quando foi pedido perdão a um grupo? Qual é o objectivo do perdão intergrupar? O perdão intergrupar pressupõe castigo adequado dos ofensores e reparação adequada por parte do grupo ofensor? Quem pode decidir se o perdão é apropriado ou não? Quem pode falar em nome do grupo? A quem devem ser dirigidas as mensagens de perdão? Qual pode ser o papel das instituições internacionais? Participaram nesse estudo 169 mulheres e 185 homens com idades oscilando entre 19 e 55 anos. O material consistiu num questionário com 82 proposições a respeito de possíveis atitudes acerca de dar perdão. Uma grande maioria dos participantes concordaram que o perdão como processo intergrupar se revestia de sentido. Menos de 10% dos participantes consideraram que não era possível a um grupo de pessoas perdoar a outro grupo de pessoas. Os participantes manifestaram ter concepções articuladas do que podia ser um processo de dar perdão interpessoal. As suas respostas estavam claramente estruturadas e emergiram sete factores interpretáveis. Os resultados observados estão de acordo com as análises e as proposições de Shriver (1995) e Govier (2002) a propósito do perdão intergrupar.

No presente estudo, como no anterior, as questões referiam-se a aspectos muito concretos do processo de pedir perdão. Estas questões inspiraram-se nos trabalhos de Tavuchis (1991), Shriver (1995) and Digeser (2001). Estes autores sugeriram, em particular, que o pedido de perdão intergrupar deveria ser um processo público conduzido em nome de toda a comunidade por alguém representativo dela que fosse especificamente mandatado para esse propósito. Exemplos de itens são os seguintes: "O pedido de perdão deve dirigir-se aos membros do governo do outro grupo", "O pedido de perdão deve acompanhar-se de novas formas de solidariedade", "Uma parte do grupo ofensor (um partido político) pode pedir perdão em nome de todo o grupo", "O pedido de perdão deve ser pronunciado durante uma cerimónia pública prevista para este efeito", "O pedido de perdão deve abranger todos os actos cometidos num determinado período de tempo", "O pedido de perdão deve ser pronunciado na língua do grupo que faz o pedi-

do”, “O pedido de perdão deve ser pronunciado a partir de uma instância de carácter mundial (a ONU, por exemplo)”. Muito embora o questionário fosse elaborado antes da publicação do livro de Amstutz (2004), muitos dos itens que foram desenvolvidos eram semelhantes aos que foram assinalados por Amstutz (2004).

Método

Participantes

Participaram neste estudo 226 pessoas (111 mulheres e 115 homens) da região de Dili. A sua idade oscilava entre 19 e 70 anos, com uma média de idade de 32 anos ($DP=10.34$). Cento e quarenta e cinco participantes completaram o ensino secundário. Cento e oitenta e quatro participantes identificaram-se como crentes em Deus, 176 declararam que costumavam perdoar na sua vida quotidiana, e 167 frequentavam a igreja (a católica) de modo habitual. Cento e trinta e cinco participantes declararam que tinham sofrido pessoalmente com a guerra, e 156 que um elemento da sua família tinha sofrido com ela. Somente 65 participantes afirmaram que não tinham sofrido pessoalmente ou através da sua família com ela.

Todos os participantes foram voluntários. Foram abordados do seguinte modo. Alguns dos participantes eram conhecidos dos experimentadores ou de um membro da sua família. Foram contactados directamente; foi-lhes explicado o objectivo do estudo, e pedido para participar. Estes primeiros participantes, após trabalharem com os experimentadores, ajudaram então a contactar com outras pessoas. O contacto directo e a confiança mútua foi necessária para convencer as pessoas a participarem no estudo.

Material

O material consistiu num questionário com 77 proposições a respeito de possíveis atitudes acerca de pedir perdão (Quadro 1). Uma escala com 17 cm foi impressa após cada proposição. Os dois extremos da escala iam de “Completamente em desacordo” a “Completamente de acordo”.

Foram utilizadas duas línguas de comunicação: o tétum e o português. O questionário foi traduzido do Português para o Tétum por um tradutor completamente bilingue, depois do Tétum para o português por um outro tradutor. Foram seguidos os passos usualmente aceites em Psicologia Intercultural (Neto, 2002).

Procedimento

Os dados foram recolhidos de Julho a Setembro de 2003. Cada participante respondeu individualmente na sua própria casa ou na universidade, dependendo do que cada pessoa julgava mais conveniente. O experimentador explicava a cada participante que tinha de ler um certo número de frases expressando alguns sentimentos ou crenças sobre pedir perdão e de avaliar o grau de acordo com o conteúdo de cada proposição.

Resultados

Cada avaliação efectuada por cada participante foi transformada num valor numérico expressando a distância (de 1 a 17) entre o ponto na escala de resposta e a origem. Estes valores foram então submetidos a análises estatísticas. Devido a um grande número de comparações, o limiar de probabilidade escolhido foi de .01.

Pode um grupo de pessoas pedir perdão a um outro grupo?

A maioria dos participantes eram claramente favoráveis à ideia de que um grupo de pessoas pode pedir perdão a um outro grupo de pessoas; a média das respostas foi de 13.65. Um *t-teste* mostrou que este valor era significativamente diferente de 9 (o valor neutro), $p < .001$ (ver também Figura 1). Cento e oitenta e cinco participantes eram favoráveis ao pedido de perdão intergrupar (pontuações superiores a 11), 22 eram neutros (pontuações de 7 a 11), e 19 não eram favoráveis a essa ideia (pontuações inferiores a 7). A idade, o género, os hábitos em perdoar, o nível de instrução e o nível pessoal ou familiar de sofrimento durante a ocupação não revelaram um impacto significativo nas respostas.

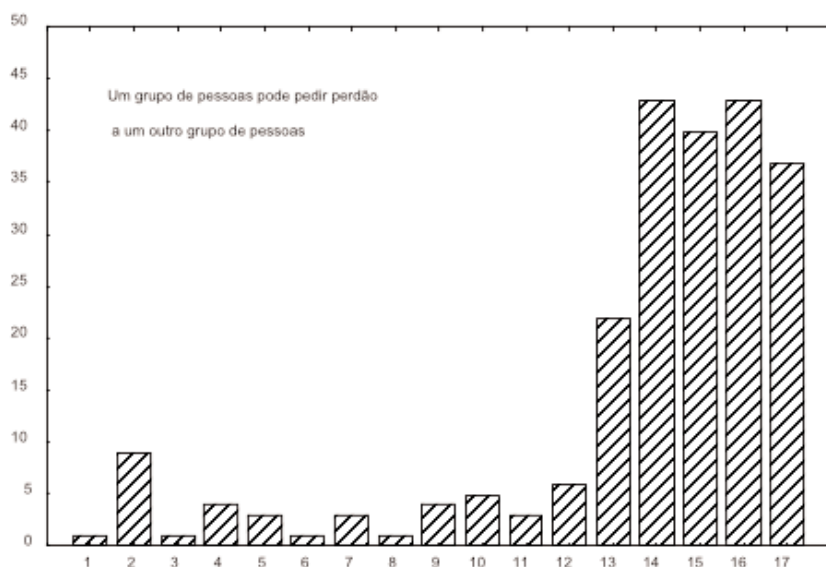


Figura 1 - *Um grupo de pessoas pode pedir perdão a um outro grupo de pessoas*

O processo de pedir perdão

Dos 76 itens restantes, 32 tinham um valor médio que era ou clara e significativamente ($p < .001$) superior (+1.50) ao valor neutro ou clara e significativamente ($p < .001$) inferior (-1.50) ao valor neutro (como mostraram os t -testes). Para os outros itens, as diferenças oscilavam entre -1.49 e +1.49, o que foi julgado demasiado próximo do valor neutro para ser considerado além disso no estudo. Foram considerados como não sendo suficientemente indicativos da presença de clara maioria das concepções acerca de pedir perdão no contexto intergrupar (ver também Quadro 1).

A principal mensagem contida nas respostas dos participantes aos 32 itens que claramente se desviavam do valor neutro era que o pedido de perdão deve ser pronunciado directamente perante uma ou mais pessoas do outro grupo (10.53), isto é, pessoas que exercem uma autoridade tradicional (11.00) ou a assembleia dos representantes do grupo a que se dirige o pedido (10.59). O pedido de perdão deve ser pronunciado por escrito (11.80) e também oralmente (11.07), na língua do grupo a que se dirige o pedido (11.93), num lugar simbólico ou sagrado do grupo a que se dirige o pedido (11.67), durante uma cerimónia pública prevista para este efeito (12.27), em nome do conjunto dos membros do grupo que faz o pedido (11.56). O pedido de perdão não deve ser feito unicamente em nome dos membros do grupo

que não são culpados de exageros (7.35). O pedido de perdão deve também ser pronunciado a partir da assembleia dos representantes do grupo que faz o pedido (11.11), deve também ser pronunciado indirectamente através da rádio ou da televisão (10.68), numa ou em várias línguas com grande difusão regional (11.18) e em várias línguas com grande difusão mundial (10.77).

As autoridades religiosas do grupo que faz o pedido podem decidir o início das diligências do pedido de perdão (10.72), mas devem ser organizadas discussões públicas na população do grupo que faz o pedido para decidir o início das diligências (11.35). É uma delegação de pessoas que deve falar em nome do grupo que faz o pedido (11.05), uma delegação composta de pessoas muito respeitadas (10.82). Pessoas de um outro grupo, alheias à ofensa, não estão bem colocadas para falar em nome do grupo que faz o pedido (7.12).

O pedido de perdão deve abranger todos os actos cometidos num determinado período de tempo (10.93) e deve acompanhar-se de actos de contrição e de arrependimento por parte dos principais responsáveis das ofensas cometidas (12.13), de actos de reparação pelas ofensas cometidas (11.59) e de propostas de renúncia a certas vantagens que poderiam comprometer a reconciliação (12.24). O pedido de perdão pode no entanto estar sujeito a condições (11.43).

O pedido de perdão deve acompanhar-se de propostas de novas formas de colaboração (12.57), de novas formas de solidariedade (12.52), de novas formas de cooperação (11.96), de novas formas de coabitação (11.57), de novas formas de aliança (11.26), de novas formas de complementaridade (11.04), de ofertas de desarmamento (11.11).

Quadro 1 - Médias e desvios-padrões para o conjunto dos itens

Itens	M	DP
O pedido de perdão deve acompanhar-se de propostas de novas formas de colaboração.	12.57	4.56
O pedido de perdão deve acompanhar-se de propostas de novas formas de solidariedade.	12.52	4.50
O pedido de perdão deve ser pronunciado durante uma cerimónia pública prevista para este efeito.	12.27	4.96
O pedido de perdão deve acompanhar-se de propostas de renúncia a certas vantagens que poderiam comprometer a reconciliação.	12.24	4.36
O pedido de perdão deve acompanhar-se de actos de contrição e de arrependimento por parte dos principais responsáveis das ofensas cometidas.	12.13	4.50
O pedido de perdão deve acompanhar-se de propostas de novas formas de cooperação.	11.96	5.14
O pedido de perdão deve ser pronunciado na língua do grupo a que se dirige o pedido.	11.93	4.95
O pedido de perdão deve ser pronunciado por escrito.	11.80	4.82
O pedido de perdão deve ser pronunciado num lugar simbólico ou sagrado do grupo a que se dirige o pedido.	11.67	4.80

Items	M	DP
O pedido de perdão deve acompanhar-se de actos de reparação pelas ofensas cometidas.	11.59	4.52
O pedido de perdão deve acompanhar-se de propostas de novas formas de coabitação.	11.57	4.75
O pedido de perdão deve ser feito em nome do conjunto dos membros do grupo que faz o pedido.	11.56	4.80
O pedido de perdão pode estar sujeito a condições.	11.43	5.02
Devem ser organizadas discussões públicas na população do grupo que faz o pedido para decidir o início das diligências.	11.35	4.88
O pedido de perdão deve acompanhar-se de propostas de novas formas de aliança.	11.26	5.15
O pedido de perdão deve ser pronunciado numa ou em várias línguas com grande difusão regional.	11.18	5.23
O pedido de perdão deve ser pronunciado a partir da assembleia dos representantes do grupo que faz o pedido.	11.11	4.82
O pedido de perdão deve acompanhar-se de ofertas de desarmamento.	11.11	5.23
O pedido de perdão deve ser pronunciado oralmente.	11.07	5.36
É uma delegação de pessoas que deve falar em nome do grupo que faz o pedido.	11.05	4.82
O pedido de perdão deve acompanhar-se de propostas de novas formas de complementaridade.	11.04	5.15
O pedido de perdão deve dirigir-se às pessoas que exercem uma autoridade tradicional.	11.00	4.59
É uma pessoa só que deve falar em nome do grupo que faz o pedido.	10.95	5.11
O pedido de perdão deve abranger todos os actos cometidos num determinado período de tempo.	10.93	4.9
Compete a certas pessoas muito respeitadas do grupo falar em nome do grupo que faz o pedido.	10.82	5.05
O pedido de perdão deve ser pronunciado numa ou em várias línguas com grande difusão mundial.	10.77	5.55
As autoridades religiosas podem decidir o início das diligências do pedido de perdão.	10.72	5.27
O pedido de perdão deve ser pronunciado indirectamente através da rádio ou da televisão.	10.68	5.50
O pedido de perdão deve ser pronunciado perante a assembleia dos representantes do grupo a que se dirige o pedido.	10.59	5.22
O pedido de perdão deve ser pronunciado directamente perante uma ou mais pessoas do outro grupo.	10.53	5.00
As autoridades tradicionais podem decidir o início das diligências do pedido de perdão.	10.43	5.12
O pedido de perdão deve ser pronunciado no território do grupo a que se dirige o pedido.	10.42	5.28
Deve fazer-se uma votação no grupo que faz o pedido para indicar quem vai levar o pedido de perdão ao outro grupo.	10.40	4.99
O pedido de perdão deve ser pronunciado a partir de um lugar simbólico ou sagrado do grupo que faz o pedido.	10.37	4.92
Deve fazer-se uma votação no grupo que faz o pedido para decidir o início das diligências do pedido de perdão.	10.35	5.11
O pedido de perdão deve ser pronunciado na língua do grupo que faz o pedido.	10.26	5.15
O pedido de perdão deve ser dirigido-se ao chefe de Estado do outro grupo.	10.25	5.31
O pedido de perdão deve ser pronunciado a partir do território do grupo que faz o pedido.	10.17	5.12

Items	M	DP
O pedido de perdão deve acompanhar-se de actos de castigo dos principais responsáveis das ofensas cometidas.	10.16	4.85
Compete às autoridades religiosas falar em nome do grupo que faz o pedido.	10.15	4.95
O pedido de perdão deve automaticamente só abranger todos os membros do grupo que faz o pedido.	10.12	5.57
O pedido de perdão deve ser pronunciado no palácio do governo do grupo a que se dirige o pedido.	10.11	5.09
O pedido de perdão deve ser pronunciado em diversos lugares do território do grupo a que se dirige o pedido.	10.09	5.22
Os homens estão mais bem colocados para falar em nome do grupo que faz o pedido.	10.09	5.15
O chefe de Estado pode decidir o início das diligências do pedido de perdão.	10.05	4.86
O pedido de perdão de um grupo em relação a outro deve ocorrer muito pouco tempo após as ofensas cometidas.	10.01	5.02
O governo pode decidir o início das diligências do pedido de perdão.	9.91	5.24
O pedido de perdão deve dirigir-se a pessoas muito respeitadas do outro grupo.	9.89	5.24
Compete às autoridades tradicionais falar em nome do grupo que faz o pedido.	9.88	5.06
O pedido de perdão deve dirigir-se às autoridades religiosas do outro grupo.	9.88	4.98
Compete aos responsáveis políticos falar em nome do grupo que faz o pedido.	9.85	5.40
O pedido de perdão deve ser feito unicamente em nome dos membros do grupo que são culpados de exageros.	9.80	5.30
O pedido de perdão deve acompanhar-se de ofertas de restituições territoriais.	9.72	5.30
O pedido de perdão de um grupo em relação a outro deve ocorrer muito pouco tempo após as ofensas cometidas.	9.68	5.01
O pedido de perdão deve dirigir-se aos membros do governo do outro grupo.	9.62	5.41
Compete ao chefe do Estado falar em nome do grupo que faz o pedido.	9.52	5.46
Compete aos membros do governo falar em nome do grupo que faz o pedido.	9.40	5.16
O pedido de perdão de um grupo em relação a outro pode intervir muito tempo após as ofensas cometidas.	9.20	5.38
O pedido de perdão deve ser feito unicamente em nome dos membros do grupo que o desejam.	9.15	5.43
O pedido de perdão deve ser pronunciado a partir de uma instância de carácter mundial (a ONU, por exemplo).	9.12	5.43
As mulheres estão mais bem colocadas para falar em nome do grupo que faz o pedido.	9.08	5.53
Uma parte do grupo ofensor (uma região, uma cidade, uma zona) pode pedir perdão em seu próprio nome.	9.03	5.44
Uma parte do grupo ofensor (um partido político) pode pedir perdão em seu próprio nome.	9.02	5.39
O pedido de perdão deve abranger alguns, mas não todos os actos cometidos num determinado período de tempo.	8.97	5.29
O pedido de perdão de um grupo em relação a outro deve ocorrer muito tempo após as ofensas cometidas.	8.82	5.06
São as pessoas oriundas de instituições mundiais que estão mais bem colocadas para falar em nome do grupo que faz o pedido.	8.79	5.16
O pedido de perdão deve ser pronunciado a partir de um outro território diferente dos dois grupos.	8.73	5.21

Items	M	DP
Uma parte do grupo ofensor (uma região, uma cidade, uma zona) pode pedir perdão em nome de todo o grupo.	8.60	5.33
O pedido de perdão deve ser pronunciado a partir de uma instância de carácter regional (OUA, por exemplo).	8.58	5.35
Uma parte do grupo ofensor (um partido político) pode pedir perdão em nome de todo o grupo.	8.48	5.27
O pedido de perdão deve ser pronunciado por uma pessoa de um outro grupo, não implicado no conflito.	8.35	5.69
O pedido de perdão deve ser pronunciado a partir do palácio do governo do grupo que faz o pedido.	8.31	5.01
O pedido de perdão deve acompanhar-se de ofertas comerciais.	8.01	5.00
O pedido de perdão deve acompanhar-se de uma dívida em dinheiro.	7.60	5.07
O pedido de perdão deve ser feito unicamente em nome dos membros do grupo que não são culpados de exageros.	7.35	5.10
São as pessoas de um outro grupo, alheias à ofensa, que estão mais bem colocadas para falar em nome do grupo que faz o pedido.	7.12	5.14

Diferenças em função do sofrimento pessoal

Foram encontradas várias diferenças significativas nas médias das respostas em função do sofrimento pessoal ($p < .001$). Os participantes que sofreram pessoalmente manifestaram um valor mais elevado que os participantes que não sofreram que o pedido de perdão deve ser pronunciado directamente perante uma ou mais pessoas do outro grupo (11.74 versus 8.79) e a assembleia dos representantes do grupo a que se dirige o pedido (11.53 versus 9.20). Os participantes que sofreram pessoalmente manifestaram um valor mais elevado que os participantes que não sofreram que o pedido de perdão deve acompanhar-se de propostas de novas formas de colaboração (14.03 versus 10.42), de novas formas de solidariedade (13.51 versus 11.03), de novas formas de cooperação (13.39 versus 9.85), de novas formas de aliança (12.22 versus 9.77), de novas formas de complementaridade (12.38 versus 9.08).

Os participantes que sofreram pessoalmente manifestaram um valor mais elevado que os participantes que não sofreram que devem ser organizadas discussões públicas na população do grupo que faz o pedido para decidir o início das diligências (12.42 versus 9.74) e o pedido de perdão pode estar sujeito a condições (12.35 versus 10.04), e o pedido de perdão deve acompanhar-se de propostas de renúncia a certas vantagens que poderiam comprometer a reconciliação (12.86 versus 11.33).

Discussão

Uma grande maioria dos participantes concordou que o perdão como um processo intergrupar faz pleno sentido. Menos de 9% dos participantes pensavam que não era possível que um grupo de pessoas pedisse perdão a outro grupo. Os resultados fundamentais eram robustos no que concerne a idade, o gênero e outras diferenças individuais. Os resultados estão em consonância com os de estudos anteriores que examinaram o perdão interpessoal da perspectiva de dar perdão (Neto, Pinto, e Mullet, 2004). Nesse estudo menos de 10% dos participantes consideravam que não era possível a um grupo de pessoas perdoar a outro grupo de pessoas.

Tal como no nosso estudo anterior, os participantes mostraram ter concepções claras do que podia ser o processo de pedir perdão intergrupar. Pedir perdão intergrupar era concebido como um processo democrático. Os participantes, muito especialmente os que sofreram pessoalmente, concordaram claramente que discussões públicas devem ocorrer antes de quaisquer ações concretas levadas a cabo por políticos, e que a assembleia de representantes para pedir ao grupo seja implicada no processo. Tal está em consonância com os resultados do nosso estudo anterior. Tal está também em consonância com a perspectiva de Tavuchnis de que um pedido de desculpa “de Muitos para Muitos” (1991, p. 98) só pode ser proferida por uma autoridade, uma pessoa verdadeiramente representativa do grupo (ver também Digeser, 2001). Para além disso, os participantes mostraram querer dar um papel especial às autoridades religiosas na ajuda para iniciar o processo de pedir perdão. Isto está em consonância, por exemplo, com o que se observou na Alemanha pouco depois do fim da Segunda Guerra Mundial (ver a Declaração de Estugarda, proveniente da Igreja Evangélica da Alemanha em 1945, e a Declaração de Darmstadt, proveniente da Igreja Luterana em 1947). Os participantes também conceberam o perdão intergrupar como um processo global e colectivo. Concordaram de modo claro que se peça perdão em nome de toda a comunidade e que tenham de se envolver todas as pessoas e todas as ações praticadas. Isto está novamente em consonância com os resultados do nosso estudo anterior.

Pedir perdão intergrupar foi concebido fundamentalmente como sendo um processo público com especial deferência em relação ao grupo ofendido. Os participantes concordaram claramente que o processo tem de ocorrer concretamente no seio das suas construções simbólicas, perante representantes do grupo ofendido, e usando a sua língua, bem como muitas outras línguas (oralmente e por escrito). Tal é consistente com os resultados do nosso estudo ante-

rior. Tal está também em consonância com a concepção de Tavuchnis do pedido de desculpas intergrupais: elas devem ser “quinta-essencialmente públicas”, e não as opiniões privadas de representantes; elas são um assunto de registo público, e os pedidos de desculpas devem ser dirigidos a uma vasta audiência tão grande quanto foi a parte ofendida e que o processo de perdão também fale às terceiras partes interessadas (Tavuchnis, 1991, p. 101, ver também Digeser, 2001).

Os participantes, em especial os que sofreram pessoalmente, viram o objectivo essencial do pedido de perdão como promovendo a reconciliação entre os dois grupos, de acordo com as perspectivas de Digeser (2001). Concor-daram que devem ser feitas concessões, se necessário, com o intuito de se facilitar o processo. Para além disso, concordaram que ambas as partes deveriam fazer planos para o futuro, sobretudo para viverem de modo mais inter-dependente e cooperativo. Isto está de acordo com os resultados do nosso estudo anterior. Este resultado também é consistente com a ideia de Tavuchnis de que pedir perdão deveria ser um prelúdio para a reconciliação entre os grupos, com os resultados de Thomas e Garrod (2002) de que os jovens da Bósnia que foram mais severamente feridos durante a guerra queriam essencialmente reconciliar-se com sérvios e croatas, e com a perspectivas de Roe et al. (1999) que a reconciliação entre as pessoas irlandesas pressupõem o conhecimento da violência passada e das humilhações (ver também Roe, 2002).

Finalmente, o perdão intergrupar foi concebido como implicando a expressão de sentimentos ou emoções particulares das pessoas que pedem perdão. O item específico, “O pedido de perdão devia ser acompanhado de actos de contrição e de arrependimento por parte das pessoas responsáveis pelas ofensas perpetradas” recebeu uma avaliação positiva. Outros itens que se referiam à oferta de várias espécies de compensações também receberam avaliações positivas. Isto está um pouco discrepante com a afirmação de Tavuchnis (1991, p. 100) que o estatuto das partes envolvidas “implica uma abordagem estilizada para a linguagem e o modo de falar que propicia pouco espaço para a espontaneidade, a flexibilidade, ou improvisações encontradas no discurso corrente”, e também é consistente com a observação de Diegeser (2001, p. 35) que o sucesso do perdão político “não depende dos estados emocionais ou internos das pessoas que perdoam.”

São de referir vários limites a esta investigação. O estudo presente foi levado a cabo somente com uma amostra de pessoas a viver em Timor-Leste que tinham sofrido recentemente com uma guerra de ocupação. Muito embora os resultados observados formem um padrão claro e compreensível, a generalização destes resultados a outras pessoas que sofreram com a guerra

com os seus vizinhos imbuí-se de incerteza. Para além disso, por razões práticas, a maior parte dos participantes eram voluntários, muitos dos quais foram contactados através de outros participantes na investigação. Muito embora tivessem sido feito esforços especiais para contactar pessoas de diferentes áreas geográficas e de diferentes níveis de instrução, não estamos seguros quanto à representatividade da nossa amostra. Será necessário levar a cabo mais investigação com amostras de outras pessoas que tenham sofrido guerras civis, tais como no Líbano, em Angola, na Guiné e em Moçambique.

Para além disso, igualmente por razões práticas, alguns aspectos do processo de pedir perdão intergrupar não foram examinados, sobretudo os conteúdos concretos das mensagens de pedido de perdão. Deveriam os pedidos pormenorizar todas ou parte das ofensas cometidas pelo grupo que pede perdão (como no discurso público do Presidente von Weisäcker em 1985 no Bundestag)? Deveriam referir-se às possíveis ofensas cometidas por outros grupos (como as recomendadas por Shriver, 1995)?

Finalmente, é necessário mais investigação sobre o processo de pedido de perdão, bem como de dar perdão, em situações complexas em que os papéis de perpetrador e de vítima não estão claramente separados, isto é, em situações em que as mesmas pessoas tenham sido, em várias proporções, simultaneamente vítimas e perpetradores.

Entretanto, é encorajante que os participantes no nosso estudo mostraram, em princípio, estar abertos à reconciliação com aqueles grupos que os fizeram sofrer: estavam interessados pelo nosso estudo e eram favoráveis a um processo de perdão intergrupar.

Referências

- Amstutz, M. R. (2004). *The healing of nations: The promise and limits of political forgiveness*. Lanham: Rowman & Littlefield.
- Azar, F., e Mullet, E. (2001). Interpersonal forgiveness among Lebanese: A six-confession study. *International Journal of Group Tensions*, 30, 161-181.
- Azar, F., e Mullet, E. (2002). Forgiveness: Overall level and factor structure in a sample of Muslim and Christian-Lebanese. *Peace and Conflict: A Peace Psychology Journal*, 8, 17-30.
- Azar, F., e Mullet, E. (2002b). Muslim and Christian in Lebanon: Common views regarding political issues. *Journal of Peace Research*, 39, 755-766.
- Azar, F., Mullet, E., e Vinsonneau, G. (1999). The propensity to forgive: Findings from Lebanon. *Journal of Peace Research*, 36, 169-181.
- Digester, P. (2001). *Political forgiveness*. Ithaca: Cornell University Press.
- Govier, T. (2002). *Forgiveness and revenge*. London: Routledge.
- Henderson, M. (1996). *The Forgiveness factor*. Salem: Grosvenor.
- Minow, M. (1998). *Between vengeance and forgiveness: Facing history after genocide and mass violence*. Boston: Beacon Press.
- Mullet, E., Girard, M., e Bakhshi, P. (2004). Conceptualizations of forgiveness. *European Psychologist*, 9, 78-86.
- Neto, F. (2002). *Psicologia Intercultural*. Lisboa: Universidade Aberta.
- Neto, F., Pinto, M. C., e Mullet, E. (2004). Perdão intergrupar: Uma perspectiva timorense. *Psicologia, Educação e Cultura*, VIII, 1, 263-288.
- Roe, M. D. (2002). Contemporary Catholic and Protestant Irish America: Social identities, forgiveness, and attitudes towards *The Troubles* of Northern Ireland. *Eire-Ireland*, 37, 153-174.
- Roe, M. D., Pegg, W., Hodges, K., e Trimm, R. A. (1999). Forgiving the other side: Social identity and ethnic memories in Northern Ireland. In J. P. Harrington and E. Mitchell (Eds.), *Politics and performance in contemporary Northern Ireland*. Amherst, MA: University of Massachusetts Press.
- Shriver, D. W. (1995). *An ethics for enemies: Forgiveness in politics*. New York: Oxford University Press.
- Smedes, L. B. (1996). *The art of forgiving*. New York: Ballantine Books.
- Tavuchis, N. (1991). *Mea Culpa: A sociology of apology and reconciliation*. Stanford: Stanford University Press.

- Thomas, J. M., e Garrod, A. (2002). Forgiveness after genocide?: Perspectives from Bosnian youth. In S. Lamb, e J. G. Murphy (Eds.) *Before forgiving: Cautionary views of forgiveness in psychotherapy* (pp. 192-211). New York: Oxford University Press.
- Worthington, E. L. (Ed.) (2005). *Handbook of forgiveness*. New York: Rutledge.

CAN A GROUP REQUEST FORGIVENESS TO ANOTHER GROUP? A TIMORESE VIEWPOINT

Félix Neto

Maria do Conceição Pinto

Universidade do Porto, Portugal

Etienne Mullet

Ecole Pratique des Hautes Etudes, France

Abstract: Can nations “engage in processes that eventuate in collective repentance and forgiveness”? (Shriver, 1995, p. 71). The present study examines this question by analyzing the views of East-Timor people who were recently involved in a war of occupation and who, for the most part, personally suffered as a result of the many conflicts. The participants were 226 people (111 women and 115 men). The material consisted of one questionnaire containing 77 sentences referring to possible attitudes about asking forgiveness. A large majority of participants agreed that forgiveness as an intergroup process makes perfect sense. Less than 9% of the participants thought that it was not possible for a group of people to ask another group for forgiveness. This basic results were robust with respect to age, gender, and other individual differences. It was consistent with the results of our previous study examining interpersonal forgiveness from the viewpoint of granting forgiveness (Neto, Pinto & Mullet, 2004).

KEY-WORDS: *Intergroup forgiveness, Timor, reconciliation.*

AValiação DA CREDIBILIDADE DE ALEGAÇÕES DE ABUSO SEXUAL: CONSENSOS E CONTROVÉRSIAS

Carla Machado

Departamento de Psicologia, Instituto de Educação e Psicologia, Universidade do Minho

Resumo

O recurso a perícias psicológicas por parte dos Tribunais para avaliação da credibilidade de menores que fazem alegações de abuso sexual é cada vez mais frequente. Esta é uma das áreas que gera mais dificuldades técnicas e éticas aos psicólogos forenses e uma das que mais polémicas tem suscitado em termos nacionais e internacionais. Neste artigo procuramos rever os aspectos desta avaliação que a literatura tem vindo a apontar como consensuais, assim como identificar e discutir as zonas de maior polémica na avaliação destes casos.

PALAVRAS-CHAVE: *Perícias, abuso sexual, credibilidade, avaliação.*

Introdução

É cada vez mais frequente o recurso dos Tribunais às perícias psicológicas em matéria de abuso sexual, usualmente com vista à clarificação de dois tipos de questões: a credibilidade das alegações de abuso formuladas pela criança e/ou seus cuidadores e o impacto destes eventuais factos na vítima. Se à última destas perguntas os técnicos têm vindo a conseguir responder, ainda que com dificuldades de ordem vária (e.g., falta de instrumentos específicos e/ou aferidos para a população portuguesa, dificuldade de avaliação dos factores moderadores do impacto, existência de crianças resilientes e de outras em que o impacto do abuso está apenas "adormecido"), a primeira questão tem suscitado dificuldades consideravelmente maiores e tem vindo a ser, tanto no plano nacional como internacional, alvo de acentuada polémica. Tais dificuldades revestem particular relevo já que, frequentemente, o testemunho da criança é a principal prova de que os juízes dispõem, uma vez que estas são situações em que frequentemente não existem outras testemunhas que possam validar o

relato da criança e que a prova médica apenas está presente numa minoria de casos (Walker, 1990).

Neste artigo procuraremos rever os aspectos desta avaliação que a literatura tem vindo a apontar como consensuais, assim como identificar e discutir as zonas de maior polémica nesta avaliação. Por último, tentaremos reflectir sobre os aspectos técnicos, profissionais e socio-políticos que têm contribuído para as dificuldades de clarificação desta matéria.

Recomendações e procedimentos usuais na avaliação da credibilidade das alegações de abuso

É possível encontrar na literatura especializada referências a diversos protocolos de avaliação da ocorrência de abuso (cf. Vizard, 1991), que diferem nas estratégias utilizadas para induzir a evocação dos factos, bem como na forma como incentivam a descrição dos pormenores dos mesmos. Há também algumas divergências no que concerne às fases do processo de avaliação propostas pelos vários autores.

Não obstante estes desencontros, a maioria dos protocolos recomenda a existência de um momento inicial da entrevista com a criança dedicado ao estabelecimento da relação, seguido de uma pré-avaliação (de alcance variável) centrada nas competências da criança relevantes para a matéria em causa (e.g., capacidade de distinguir verdade de mentira, compreensão do motivo da avaliação, compreensão da noção de toque, identificação das várias partes do corpo e nomes para as designar).

Segue-se a fase de questionamento destinado à eliciação da narrativa de abuso. É também consensual o facto de esta dever começar por questões de natureza geral (e.g., "Que tipo de pessoa é o teu pai/a tua mãe?"), prosseguindo para questões cada vez mais focalizadas (e.g. "Por vezes, há coisas que acontecem que nos deixam tristes ou envergonhados. Já te aconteceu alguma coisa desse tipo?"). Caso tenha havido uma revelação anterior ou denúncia por parte do cuidador da criança, podem ser feitas questões referentes a esta situação (e.g., "A tua professora quis que tu falasses comigo por causa do que tu lhe contaste há alguns dias, sobre o que o teu primo te tem andado a fazer. Para eu te ajudar, preciso que me contes o que tem acontecido.").

A generalidade dos protocolos compreende ainda a possibilidade de se fazerem perguntas mais directas sobre toques inadequados, salientando, no entanto, que estas deverão ser formuladas de forma aberta e não sugestiva (e.g., "Já alguma vez alguém te tocou de uma forma que te fizesse sentir mal?

Que fosse segredo?"; "Há alguma parte do teu corpo em que as outras pessoas não devem tocar? Isso alguma vez te aconteceu?"). É também usualmente permitido formular questões específicas sobre toques na região genital ou anal, desde que igualmente não indutoras e colocadas de modo a que a criança as entenda (e.g., "Já alguém te tocou no teu 'pipi' sem ser para te limpar ou lavar?").

A adequada formulação das perguntas e o seu carácter não indutor são de extrema importância, dadas as questões jurídicas implicadas nestes casos. Sendo assim, é necessário adoptar particulares cuidados (Hewitt, 1999; Vizard, 1991 cf. Machado e Antunes, no prelo):

- a) evitar repetir perguntas e respostas (tal poderá ser interpretado pela criança como sinal de que a sua resposta anterior estava errada),
- b) evitar o uso do "porquê" (pode ser sentido como uma crítica),
- c) adaptar as questões ao nível desenvolvimental da criança,
- d) evitar as perguntas formuladas no formato sim ou não,
- e) evitar as perguntas de escolha múltipla (se necessárias, voltar a colocá-las mais tarde, com as opções numa ordem diferente),
- f) não fazer perguntas que contêm múltiplas interrogações,
- g) evitar usar a voz passiva ou a negativa,
- h) fornecer às crianças mais novas um enquadramento narrativo e uma estrutura cronológica para o relato ("fala-me agora de... e depois?"),
- i) evitar questões abstractas (as crianças, sobretudo as mais novas, focam-se em acções e acontecimentos e não em sentimentos ou intenções),
- j) não dirigir questões sobre os afectos ou intenções dos outros a crianças em idade pré-escolar.

Após este questionamento, e na eventualidade de a criança narrar o episódio abusivo, é também consensual que a entrevista deverá finalizar com a recapitulação dos factos pelo entrevistador, pedindo à criança que preste atenção de forma a corrigir eventuais erros. Como afirmam Machado e Antunes (no prelo), "deverá ainda elogiar-se a criança pela sua coragem em contar o sucedido, respondendo às suas eventuais questões e baixando o tom emocional da sessão, por exemplo, retomando com a criança, se esta assim o desejar, uma actividade de natureza mais lúdica".

Um bom exemplo - entre outros - deste modelo de entrevista é o protocolo CARES, do Hospital de St. Lukes (Boise, Idaho), que a investigação empírica (Cantlon, Payne e Erbaugh, 1996 cit. Hewitt, 1999) tem demonstrado ser eficaz com crianças de idades superiores a 4 anos. Também a entrevista Step-Wise (Yuille, 1996 cit. Hewitt, 1999), cujas orientações estão subjacente ao

Memorandum of Good Practice (1992) recomendado como procedimento-tipo na Grã-Bretanha, obedece à generalidade destes princípios.

Posteriormente ao momento de produção da narrativa de eventual abuso e no que diz respeito à ponderação das declarações produzidas e emissão de juízo sobre a sua credibilidade há também algumas zonas de acordo entre autores:

- a) a concepção do juízo de credibilidade como uma avaliação geral, baseada no conjunto dos indicadores recolhidos, e não equiparável ao resultado de uma qualquer prova/instrumento de avaliação (Ney, 1995);
- b) a necessidade de, para emitir este juízo, confrontar o relato dos factos feito pela criança com as suas características individuais (e.g., inteligência, personalidade, nível desenvolvimental);
- c) a necessidade de considerar o perfil sintomatológico da criança, ainda que estando atento, quer à possibilidade de simulação, quer à existência de crianças abusadas que são resilientes, de facto ou pelo menos a curto prazo (Saywitz, Mannarino, Berliner e Cohen, 2000);
- d) a ponderação cautelosa das circunstâncias em que a alegação de abuso foi feita (por quem, quando, como, com que objectivos) e de que forma o processo de acusação/investigação se foi desenrolando no tempo (Ehrenberg e Elterman, 1995; Bow, Quinnell, Zaroff e Assemany, 2002);
- e) a necessidade de considerar a possibilidade de hipóteses alternativas explicativas das alegações (e.g., fantasia, mentira, sugestão, erro de identificação) antes de emitir um juízo final (e.g., Yuille, Tymofievich e Marxsen, 1995; Ney, 1995; Ehrenberg e Elterman, 1995; Bow et al., 2002);
- f) a emissão do juízo do avaliador em termos probabilísticos (e.g., "abuso provavelmente falso"; "declarações compatíveis com uma alegação de abuso"), compreendendo a possibilidade de juízos inconclusivos (e.g., Yuille, Tymofievich e Marxsen, 1995; Ehrenberg e Elterman, 1995).

Um consenso relativo também existe em questões de formação e treino, quer no que diz respeito à necessidade de formação específica, por exemplo para usar bonecas anatomicamente correctas, quer na maior importância relativa deste treino quando comparado com outras variáveis, como sejam o sexo do avaliador (Vizard, 1991).

Finalmente, no plano ético, há também um consenso alargado em algumas matérias, tais como o dever de neutralidade, a clareza quanto ao pagamento da avaliação (estabelecendo previamente o montante a pagar e clarificando que este se refere ao trabalho produzido e não à opinião a emitir) ou a separação entre o processo de avaliação e tratamento (o terapeuta não deverá prestar declarações num processo como perito forense sobre o seu cliente, embora o possa fazer como clínico¹).

Brevemente descritas estas áreas de relativo consenso, quais são então as matérias que têm contribuído para tornar este um dos mais polémicos e mediáticos temas da psicologia?

Polémicas e controvérsias

Da análise da literatura resulta claro que as zonas de maior conflitualidade em matéria de avaliação de crianças alegadamente vítimas de abuso dizem respeito à utilização de instrumentos auxiliares da revelação (com particular relevo para as bonecas anatomicamente correctas, mas também a propósito do uso de desenhos), assim como o recurso a indicadores de veracidade sugeridos por alguns autores. Menos centrais na discussão pública sobre esta matéria, mas também pouco consensuais, são o número de entrevistas necessárias para proceder à avaliação, o número de entrevistadores que deve estar presente na sala, a quantidade de informação prévia de que o técnico deve dispor antes de analisar o caso, a gravação electrónica das entrevistas e o valor diagnóstico de alguns instrumentos de avaliação, quer os testes psicológicos de largo espectro, quer os instrumentos centrados na sintomatologia supostamente apresentada pelas crianças abusadas.

Analisemos, pois, cada uma destas matérias.

a) Bonecas anatomicamente correctas

A área de maior polémica no que diz respeito à avaliação do abuso diz, indubitavelmente, respeito à utilização de bonecas anatomicamente correctas, isto é, de bonecas com órgãos sexuais, pilosidade e orifícios na região oral e anal. Apesar de já usadas anteriormente no contexto hospitalar, nomeadamente na explicação às crianças de algumas intervenções médicas, é apenas

¹ Menos consenso existe sobre se o perito poderá posteriormente, findo o processo legal, aceitar aquele cliente em terapia. A APA (1998), por exemplo, desencoraja esta prática e recomenda que, se tal for feito, sejam assumidos particulares cuidados.

em 1976 que estas começam a ser aplicadas no âmbito da avaliação do abuso (Morgan, 1995).

Particularmente importante é a noção de que o seu uso pode ter múltiplas finalidades, mas que estas não constituem um teste (Hewitt, 1999; Koocher, Goodman, White, Friedrich, Sivan e Reynolds, 1995). Segundo Everson e Boat (1994 cit. Koocher et al., 1995) as bonecas podem ser usadas com seis finalidades diferentes: a) para criar um ambiente menos ameaçador durante a entrevista; b) para introduzir o tema sexual e transmitir autorização para o abordar, c) para facilitar a identificação das diferentes partes do corpo; d) como um auxiliar da demonstração do que sucedeu; e) como um estímulo mnésico; e f) como um auxiliar da revelação. Em nenhum destes casos, os jogos sexuais que as crianças podem fazer com as bonecas (e.g., colocar os dedos no orifício anal, explorar os órgãos genitais) deverão ser considerados evidências de abuso (Ney, 1995), "sendo antes particularmente importante aquilo que a criança revela verbalmente ou demonstra com as bonecas, a partir das perguntas do avaliador" (Machado, 2002, p.59).

De que forma podem ser usadas e apresentadas as bonecas?

Na generalidade dos protocolos as bonecas são apresentadas à criança vestidas (cf. Vizard, 1991), ainda que haja autores que consideram que o contrário deve ser feito (Boat e Everson, 1986 cit Koocher et al., 1995) e outros que consideram que a questão é irrelevante (Friedemann e Morgan, 1985 cit Koocher et al., 1995). A maioria dos autores sugere o uso de várias bonecas, dos dois sexos e de diferentes idades, e a grande maioria das directrizes publicadas sobre o tema sugere que, num primeiro momento, as questões do avaliador se devem centrar na identificação das diferentes partes do corpo e das designações da criança para as mesmas (Koocher et al., 1995).

Mais controversa é a questão do emparelhamento étnico das bonecas com a vítima, já que, apesar de haver algumas recomendações neste sentido (e.g., Boat e Everson, 1986 cit Koocher et al., 1995), não há qualquer evidência empírica de que tal seja uma prática facilitadora. Como advertem Koocher e colaboradores (1995), a criança pode ter antecedentes étnicos mistos ou pode ter sido abusada por mais do que um indivíduo de grupos étnicos diferentes. Mais importante ainda, não há qualquer evidência de que a pista racial seja importante na selecção que a criança faz de um boneco para representar o abusador (e.g., a cor do cabelo, forma do rosto ou dos olhos podem ser igualmente relevantes para a escolha daquele boneco específico).

A área central da polémica sobre o uso das bonecas reside, contudo, no momento em que as mesmas devem ser introduzidas, intimamente relacionado com o fim para que são usadas. Alguns autores sugerem que estas apenas de-

vem ser introduzidas depois de a criança ter feito alguma referência ao abuso (Morgan, 1995) e outros apresentam as bonecas após a fase inicial da entrevista, apenas colocando questões mais específicas sobre toques indesejados e genitais após a sua introdução (e.g., White, 1987 cit. Vizard, 1991). Assim, enquanto no primeiro procedimento as bonecas funcionam essencialmente como um auxiliar da descrição dos factos, no segundo caso as bonecas têm, para além desta função, o objectivo de promover a revelação (na medida em que a criança recebe a mensagem implícita de que é possível falar sobre temas e órgãos sexuais). Obviamente que o primeiro destes modos de procedimento suscita menos dúvidas e oposição, na medida em que minimiza a possibilidade de indução do avaliador.

Independentemente desta escolha do momento e finalidade do uso das bonecas, já que este procedimento faz apelo à capacidade de simbolização, tem sido recomendado que aquelas não sejam usadas com crianças abaixo dos três anos e meio/quatro, uma vez que apenas por volta desta idade está consolidada a capacidade de representação do *self* (Hewitt, 1999). Alguns estudos (e.g., Bruck, Ceci, Francoeur e Renick, 1995) têm verificado que o uso das bonecas em idades mais precoces pode conduzir a erros interpretativos.

Não obstante estes cuidados, a crítica central apontada à utilização de bonecas anatomicamente correctas nas avaliações de abuso reside na sua suposta indução de falsas revelações, pela presença de órgãos sexuais (cf. Vizard, 1991). A investigação sobre esta alegação (cf. Hewitt, 1999; Vizard, 1991; Davies, 1991) não tem, contudo, validado este receio, afirmando que as bonecas não aumentam o número de falsas revelações. A excepção parece encontrar-se precisamente no caso das crianças mais novas (e.g., 3 anos), mais susceptíveis de produzir falsas revelações quando as bonecas são usadas (Bruck et al., 1995). Apesar de o limiar usual de uso ser estabelecido por volta do 4 anos, uma investigação recente de Bruck, Ceci e Francoeur (2000) verificou que este problema se mantém até aos 5 anos, recomendando que as bonecas não sejam usadas com idades inferiores a este limiar etário.

A questão das falsas revelações (isto é, referências verbais a actos abusivos) deve ser separada da exploração das bonecas e de eventuais toques sexualizados com estas. Uma revisão dos estudos feitos com bonecas anatomicamente correctas (Koocher et al., 1995) demonstrou, de forma bastante consensual, que as crianças de amostras normativas têm frequentemente interesse e tocam nos órgãos sexuais das bonecas. Por outro lado, nestes grupos normativos, era raro haver demonstrações explícitas de actos sexuais ou manifestações intensas de agressividade (ibidem).

Comparativamente, tem sido dito que as crianças abusadas têm com as bonecas um número significativamente maior de jogos e interações sexualizadas (cf. Vizard, 1991; Hewitt, 1999), uma conclusão que tende a ser sustentada pela revisão da investigação efectuada por Koocher e colaboradores (1995), ainda que haja estudos em que não se encontraram quaisquer diferenças. Para além de estes comportamentos, como salientámos anteriormente, não deverem ser tomados como evidências lineares de abuso, há a recordar que os comportamentos sexualizados podem ser estimulados por outras características adversas do ambiente familiar. Por sua vez, August e Forman (1989 cit. Koocher et al., 1995) sugerem que pode haver diferentes tipos de reacções às bonecas entre as crianças abusadas, desde aquelas que exibem um comportamento abertamente agressivo e sexualizado, até às que adoptam uma postura de evitamento da interacção com as bonecas.

Como conclusão desta revisão, Koocher e colaboradores (1995) afirmam que "usar as bonecas anatomicamente correctas nas avaliações claramente não é em si mesmo perturbador ou hiper-estimulante para as crianças" (p. 211). Acrescentam que "usar as bonecas pode assistir na identificação dos nomes preferidos ou idiossincrásicos que as crianças usam para as diferentes partes do corpo" (ibidem) e que a mensagem geral produzida pela investigação é consistente com a afirmação da APA de que "tais bonecas podem ser úteis para ajudar as crianças a comunicar quando as suas competências linguísticas ou preocupações emocionais impedem respostas verbais directas" (American Psychological Association, 1991 cit. Koocher et al., 1995, p.211).

Uma outra polémica, relacionada com a eventual interacção entre o recurso às bonecas e a menor capacidade de as crianças resistirem a questões sugestivas, não tem também encontrado muito suporte na investigação empírica, ainda que esta seja esparsa (cf. Davies, 1991). Há, contudo, alguns estudos que sugerem que o recurso às bonecas poderá aumentar a margem de erro em algumas crianças se forem usadas questões inadequadas (Koocher et al., 1995). Finalmente, ainda que seja necessário cuidado na distinção entre fantasia e realidade, não há evidências consistentes de que as crianças pequenas tenham fantasias sexuais explícitas ou de que as confundam com factos reais (Koocher et al., 1995), algo que por vezes é alegado em relação às revelações de abuso na presença de bonecas.

Assim, dos dez protocolos reconhecidos na prática britânica no início dos anos noventa (cf. Vizard, 1991), sete incluíam o uso rotineiro das bonecas, um admitia tal possibilidade, um outro era omissivo sobre o assunto e um único protocolo recusava explicitamente o recurso a este instrumento. O *Memoran-*

dum of Good Practice (1992), recomendado como orientador pelo Home Office, é também omissa quanto a este aspecto, ainda que Yuille, um dos autores da entrevista Step-Wise na qual o *Memorandum* se baseia, seja desfavorável ao uso das bonecas (cf. Vizard, 1991).

As orientações publicadas pela *American Psychological Association* e pela *American Professional Society on the Abuse of Children* (1995, 1997 cit. Hewitt, 1999) também admitem o uso das bonecas como auxiliares da demonstração ou estímulos para ajudar as crianças a mostrar o sucedido. Em síntese, Koocher e colaboradores (1995) afirmam que "a investigação até à data maioritariamente apoia o uso de bonecas anatomicamente correctas como auxiliares da comunicação ou da memória para crianças com 5 anos ou mais, ainda que com um certo risco de contribuir para os erros de algumas crianças se forem usadas questões sugestivas" (p.217).

b) Desenhos anatómicos

A par das bonecas, os desenhos anatomicamente correctos integram grande parte dos protocolos utilizados internacionalmente, nomeadamente o protocolo CARES (St. Lukes Hospital, Idaho), a entrevista Step-Wise (Yuille, 1996 cit. Hewitt, 1999) ou o Touch Survey (Hewitt, 1999). No protocolo CARES o entrevistador desenha um corpo humano e questiona a criança sobre partes que não devem ser tocadas e sobre se alguém já lhe tocou nos órgãos genitais, enquanto na entrevista Step-Wise o desenho da figura humana é utilizado para identificar e nomear as diferentes partes do corpo, seguida de questões sobre toques na zona genital.

O Touch Survey constitui uma proposta de avaliação também baseada no desenho, dirigida a crianças em idade pré-escolar. Neste procedimento, com recurso a desenhos muito simples que o avaliador vai fazendo, a criança é questionada sobre vários tipos de toques (positivos, negativos e neutros), e sobre em que parte do corpo estes ocorreram. Questões sobre quem foram as pessoas que os tocaram dessa forma e quais os sentimentos elicitados pelo toque são também formuladas. Caso haja referência a toques que possam configurar uma situação de abuso, poderá recorrer-se às bonecas anatomicamente correctas para demonstração da interacção física ocorrida. Apesar do interesse deste procedimento e do seu usual bom acolhimento pelas crianças, é de salientar que o mesmo não foi ainda empiricamente validado.

Entre nós, o manual CORE, adoptado pela Associação Portuguesa de Apoio à Vítima, recorre também ao uso de desenhos anatómicos. Não conhecemos, contudo, qualquer avaliação da eficácia deste procedimento.

Ainda que com as reservas assinaladas, o recurso a desenhos anatomicamente correctos não deve ser confundido com o recurso à interpretação do desenho infantil como estratégia de despistagem do abuso. Isto é, há que diferenciar o uso de desenhos para questionar a criança ou ajudá-la a descrever os eventos abusivos, da inferência de abuso efectuada a partir de pistas supostamente contidas no desenho da criança. Ainda que existam algumas propostas neste sentido (e.g., "Favourite Kind of Day Drawing", Manning, 1987 cit. Veltman e Browne, 2002), uma recente revisão dos estudos empíricos que procuram avaliar a credibilidade deste tipo de inferências levanta sérias dúvidas sobre a mesma. Assim, os autores concluem que, ainda que os desenhos espontaneamente produzidos que contêm órgãos sexuais "possam requerer uma investigação suplementar, no presente parece haver a possibilidade de o uso isolado deste indicador produzir um alto nível de falsa identificação de maltrato sexual" (Veltman e Browne, 2002, p.34).

c) Indicadores de credibilidade

Várias têm sido as tentativas para encontrar indicadores empíricos que possam ajudar a determinar a credibilidade das alegações de abuso, baseadas na premissa de Undeutsch (1989 cit. Hewitt, 1999) de que as pessoas que experienciaram um dado evento o relatarão de forma diferente das que apenas tiveram acesso indirecto ao mesmo.

Podemos citar, a título de exemplo, a investigação conduzida por Faller (1988 cit. Davies, 1991), que destaca como características típicas dos relatos verdadeiros de abuso a referência a detalhes sexuais explícitos, a resposta emocional intensa da criança e a sua capacidade de identificar as circunstâncias e o local do evento. Estes aspectos estão também presentes nos indicadores de veracidade propostos por Salter (1988 cit. Burkhardt e Rotatori, 1995), que acrescenta ainda a relevância diagnóstica de detalhes ou actos descritos de uma perspectiva especificamente infantil (e.g., "fiquei molhada porque ele fez chichi nas minhas pernas"), dos relatos que incluem pormenores relacionados com experiências sensoriais (odor, textura, calor, sabor) e ainda a presença no relato de detalhes típicos deste tipo de ofensas (instruções de segredo, ameaças, escalada de actos abusivos).

Curiosamente, todos estes indicadores fazem parte do mais conhecido protocolo de avaliação da veracidade do abuso, o *Statement Validity Analysis* (Kohnken e Steller, 1988 cit. Vrij, 2005). Este protocolo consiste numa adaptação e melhoria de um procedimento de análise das declarações, usado desde a década de 50 nos Tribunais alemães para lidar com casos em que a peça

central de evidência era o testemunho não corroborado de crianças. A análise da credibilidade do testemunho com base no SVA é actualmente aceite como evidência em vários países europeus, nomeadamente a Alemanha, a Suécia e a Holanda.

A partir dos critérios iniciais de discriminação da veracidade propostos por Undeutsch (1967 cit. Vrij, 2005), vários autores, entre os quais os mais conhecidos são Steller e Kohnken, desenvolveram sucessivos esforços de aperfeiçoamento desta metodologia. O ponto de chegada foi o *Statement Validity Analysis*, um procedimento composto por três fases:

- a) uma entrevista semi-estruturada, não indutiva, que é gravada e transcrita;
- b) a análise de conteúdo do relato, centrada na sua validade interna, e
- c) a aplicação de uma *checklist* de validade, que confronta o relato com indicadores externos.

A mais conhecida destas fases é, indubitavelmente, a segunda, frequentemente confundida aliás com o procedimento no seu todo. Esta análise de conteúdo, usualmente designada por CBCA (*Content-Based Criteria Analysis*) propõe que cada testemunho seja avaliado em dezanove aspectos, agrupados em cinco grandes dimensões: estrutura formal (e.g., consistência interna, detalhe), aspectos estruturais (e.g., menção a circunstâncias de tempo e lugar, a verbalizações do ofensor), especificidades do conteúdo (e.g., detalhes sexuais inusuais, interpretações infantis), estado motivacional (e.g., correcções espontâneas) e sinais comuns no abuso (e.g., escalada nos casos de incesto).

No plano teórico, Kohnken (cf. Vrij, 2005) considera que estes indicadores se distribuem entre o plano cognitivo - características da narrativa que indicam veracidade porque são de difícil "fabricação" - e motivacional - elementos tipicamente presentes nas histórias verdadeiras mas que contrariam o estereótipo da verdade (e.g., admitir falhas de memórias ou dúvidas sobre o próprio relato).

Independentemente do grau de validade do CBCA, que analisaremos mais adiante, a investigação tem demonstrado claramente que vários factores podem interferir com o desempenho da criança nestes critérios: "o facto de a criança ter sido obrigada a repetir o sucedido a diversos entrevistadores (fazendo diminuir a tonalidade emocional da narrativa e a sua espontaneidade), a natureza breve ou superficial do contacto (implicando, obviamente, menos detalhes sobre os factos) ou pelo contrário, a sua natureza repetida e constante (que pode levar à dificuldade em isolar eventos passível de relatar com pormenor) e ainda o grau de violência do abuso (eventos percebidos como muito

ameaçadores pela criança podem conduzir à omissão e mesmo à dissociação dos detalhes mais traumáticos como uma forma de protecção contra o trauma" (cf. Machado e Antunes, no prelo). Particular cautela merece a aplicação deste procedimento a crianças de idade pré-escolar, que alguns autores (e.g., Hewitt, 1999) desaconselham. Efectivamente, o nível desenvolvimental destas é incompatível com alguns dos indicadores de veracidade exigidos (e.g., nível de detalhe, relato espontâneo sem necessidade de estruturação pelo avaliador, expressão espontânea de afecto, referência a intenções ou motivos do abusador...) e o uso do CBCA nestes casos poderá conduzir a falsos negativos.

O terceiro componente do SVA, a *checklist* de validade, foi desenvolvido precisamente para tentar lidar com estes factores que podem afectar a pontuação no CBCA. Apesar de existirem algumas variações nos factores que integram esta *checklist*, esta inclui elementos como a capacidade expressiva da criança, a sua vulnerabilidade à sugestão, a adequação no afecto com que os factos são relatados, o contexto em que a revelação aconteceu, a adequação da forma como a entrevista foi conduzida, as pressões externas a que a criança pode estar sujeita, entre outros.

O que nos diz a investigação, composta por estudos de terreno e experimentais (na maioria dos casos apenas usando o CBCA), sobre a validade deste procedimento?

Há autores que consideram que o nível de precisão atingido com estes procedimentos é claramente insuficiente para a utilização do CBCA como um teste fiável e válido que deva ser utilizado em contexto forense (Lamb et al., 1997 cit Hewitt, 1999), a par de outros (e.g., Arce e Fariña, 2005; Yuille, Tymofievich e Marxsen, 1995; Honts, 1994 cit. Ruby e Brigham, 1997) que consideram que o seu nível de eficácia é elevado e que é um procedimento válido. Arce e Fariña (2005), por exemplo, referem uma taxa de classificação acertada em contexto laboratorial que varia entre 65% e 85% dos relatos e níveis de eficácia que chegam aos 100% em estudos conduzidos com casos reais. Uma posição mais cautelosa é assumida por autores como Horowitz, Lamb, Esplin, Boychuk, Krispin e Reiter-Lavery (1997) que consideram que, "com alguma revisão, o sistema CBCA tem suficiente fidelidade para ser usado como um procedimento psicométrico, desde que adequada validade possa também ser demonstrada" (p.11).

Em oposição, num estudo de Lamer-Winkleman, Buffing e van der Zanden (cit. Bekerian e Dennett, 1995) o CBCA não conseguiu discriminar entre casos substanciados e infundados. Outros estudos (cf. Bekerian e Dennett, 1995) têm também divergido quanto aos resultados encontrados, destacando a necessidade de mais investigação sobre este procedimento.

A avaliação geral da (pouca) investigação conduzida até ao momento, feita por alguns autores de referência (Davies, 1991; Yuille, Tymofievich e Marxsen, 1995; Vrij, 2005; Ruby e Brigham, 1997), parece apontar este protocolo como globalmente útil para a diferenciação entre verdadeiras e falsas alegações. Isto é, a maioria dos estudos encontra de facto diferenças entre a pontuação que os relatos falsos e verdadeiros recebem no CBCA, diferenças estas que vão na direcção esperada. Na sua recente revisão dos estudos feitos sobre este método, Vrij (2005) conclui que a taxa média de classificações acertadas com recurso ao CBCA é de 73%. Estes resultados são francamente melhores do que os obtidos, por exemplo, pelo recurso a indicadores não-verbais de mentira (Vrij, 2005). Algumas evidências têm também sido encontradas no sentido de sustentar que o CBCA poderá também ser útil com adultos e na avaliação de outras situações que não só o abuso sexual (Vrij, 2005; Ruby e Brigham, 1997).

Apesar destes resultados positivos, há várias reservas a apontar ao método, a primeira das quais é a percentagem, apesar de tudo elevada (cerca de 30%) de erros que este permite.

Problemas existem também no plano procedimental, existindo divergências sobre a forma de pontuação dos itens da CBCA (alguns autores analisam somente a sua presença vs. ausência enquanto outros sugerem a sua cotação numa escala de tipo Likert), assim como quanto ao número de indicadores que devem estar presentes para que um relato seja considerado verdadeiro. Embora algumas sugestões de regras de decisão tenham sido avançadas (cf. Ruby e Brigham, 1997), a maioria dos peritos opõe-se a estas, uma vez que vários factores podem afectar o desempenho em itens específicos do CBCA, tal como atrás referimos (Vrij, 2005).

Questões existem também quanto à ponderação de cada item, havendo sugestões de que nem todos têm o mesmo valor diagnóstico (Bekerian e Dennett, 1995). O item 3 (quantidade de detalhes) é o que se tem revelado mais robusto, havendo também resultados muito sólidos quanto a indicadores como a produção não estruturada, descrição do contexto e capacidade de reproduzir as conversas com o ofensor (Ruby e Brigham, 1997; Vrij, 2005). Pelo contrário um dos critérios da lista (auto-depreciação) não recebeu até agora qualquer validação empírica (ibidem). No seu conjunto, os indicadores cognitivos parecem mais válidos do que os indicadores motivacionais (Vrij, 2005). Esta falta de uniformidade, assim como os resultados obtidos nos estudos que analisam a consistência entre vários codificadores, têm conduzido à conclusão de que é mais fiável usar o score total do CBCA do que recorrer a itens isolados (Vrij, 2005; Ruby e Brigham, 1997; Horowitz et. al., 1997).

Finalmente, há a considerar os factores que podem afectar a pontuação obtida na CBCA. Para além da idade, tem sido verificado que esta pode ser alterada pelo estilo da entrevista (quanto mais sugestiva, maior a pontuação), pelo treino prévio dos sujeitos, pela exposição da criança a sugestões de adultos informados sobre o método, e por características pessoais da criança, tais como a sua ansiedade social (Vrij, 2005; Vrij, Akehurst, Soukara e Bull, 2004). Particularmente complexas, a nosso ver - mas não só difíceis de avaliar com o CBCA, como com qualquer outro procedimento - são os casos de falsa identificação (em que o relato é verdadeiro mas o agressor é outro que não o apontado) e as situações de falsas memórias, em que a criança acredita de facto ter passado pela experiência narrada. Alguns autores consideram que, para além dos conhecidos casos de falsas memórias induzidas terapeuticamente, poderão ocorrer, provavelmente com muito maior frequência, situações em que eventos repetidamente descritos à criança são erradamente identificados por esta (sobretudo se for muito nova) como tendo sido directamente experienciados (Ruby e Brigham, 1997). O único estudo que procurou, com o nosso conhecimento, analisar a eficácia do CBCA na discriminação destas pseudomemórias de relatos verdadeiros (Huffman e Ceci, 1997 cit. Ruby e Brigham, 1997) verificou que estes últimos obtinham em média pontuações mais altas no CBCA, mas que era grande a margem de sobreposição das duas amostras.

Ainda que a terceira fase do método, a *checklist* de validade, tenha procurado lidar com alguns destes factores, a investigação mostra que esta por vezes não é usada ou que o é de forma menos sistemática e que a sua aplicação é bastante mais subjectiva do que a do CBCA. Dúvidas teóricas têm também sido levantadas (Vrij, 2005) sobre a adequação de alguns dos itens que compõem esta *checklist*. Desta forma, o procedimento muitas vezes acaba por se confundir com a CBCA, o que aliás tem acontecido também na investigação.

A análise cuidadosa dos resultados da investigação revela, pois, um quadro complexo, que aponta o SVA como um instrumento promissor e com potencial utilidade, mas salienta, ao mesmo tempo, várias dificuldades e potenciais erros do mesmo. Compreende-se, pois, que os autores que se têm dedicado à revisão dos estudos sobre este procedimento, tais como Vrij (2005) e Ruby e Brigham (1997), assumam uma posição cautelosa, não recomendando o uso do SVA como um teste em contexto forense mas encorajando a investigação sobre o mesmo e apontando o seu potencial uso como um "indicador grosseiro da veracidade" (Vrij, 2005, p.34) de um relato, que deverá ser sujeito a um escrutínio posterior mais cauteloso. Do nosso ponto de vista e

no protocolo de avaliação que construímos (cf. Machado e Antunes, no prelo), os indicadores de veracidade deverão ser utilizados com cuidado e integrados com outros elementos na avaliação global do caso.

d) Número de entrevistas e de entrevistadores

Não há acordo quanto a esta matéria, com sugestões que variam entre uma única sessão e um número indeterminado destas (Vizard, 1991). Embora vários protocolos pareçam envolver uma única sessão, outros apresentam uma estrutura mais flexível (e.g., Machado e Antunes, no prelo) e alguns autores (Arce e Fariña, 2005) salientam a necessidade de mais do que uma sessão de forma a verificar a consistência temporal do relato da criança. A necessidade de mais do que uma sessão com cada um dos diferentes elementos envolvidos na avaliação tem também sido justificada pela quantidade de informação que é necessário recolher (Bow et al., 2002), nomeadamente a história psicossocial da criança e da família, avaliação desenvolvimental, respostas à revelação, etc. Assim, a necessidade de mais do que uma entrevista provavelmente manter-se-á, ainda que muitas crianças revelem o abuso logo na primeira ou segunda sessão (London, Bruck, Ceci e Shuman, 2005).

Por sua vez, no que diz respeito ao número de avaliadores, alguns autores sugerem que poderá ser importante a presença de dois técnicos na sala, deixando um mais disponível para interagir com a criança e outro focalizado no registo do relato e na observação do comportamento não verbal da criança (Barker, 1990). Por outro lado, é verdade que a presença de outra pessoa pode reforçar a inibição já frequente nestes casos (ibidem) e alguns autores consideram mesmo que a existência de múltiplos entrevistadores pode comprometer a capacidade de testemunhar e ser psicologicamente danosa para a criança. A gravação em vídeo ou o recurso (com informação à criança) ao espelho unidireccional poderão ser alternativas mais adequadas (Barker, 1990; Saywitz, Goodman e Lyon, 2002).

e) Informação prévia

Este é também um tópico que tem merecido considerável disputa, havendo autores (Great Ormond Street Team cf. Vizard, 1991; Bow et al., 2002) que salientam o valor da informação prévia recolhida pelo avaliador (e.g., declarações já recolhidas, peritagens médicas) e outros que propõem um modelo de avaliação "cego" (protocolo CARES, cf. Hewitt, 1999). Ainda que exista alguma evidência sobre a eficácia da avaliação cega quando comparada com situações de avaliação "informada" (ibidem), não é inteiramente claro se tal

se deve especificamente a esta variável ou ao formato mais geral de avaliação adoptado no âmbito do protocolo CARES. Por outro lado, concordamos com Barker (1990) quando este destaca a importância prática da recolha prévia da máxima informação possível pelo avaliador, quer no sentido de facilitar a abordagem do evento (por exemplo, se tiver havido previamente revelação, tal pode ser usado como um ponto de partida para a entrevista), quer para evitar factores de contaminação da avaliação (por exemplo, entrevistar a criança sendo esta trazida à consulta pelo presumível abusador). Bow e colaboradores (2002) destacam também a importância de recolher informação sobre os interrogatórios prévios a que a criança foi submetida e seu contexto, de forma a minimizar o possível viés decorrente das inquirições sucessivas, frequentes nestes casos.

f) Gravação das entrevistas

A necessidade de ser possível auditar o processo e a exigência, em certos protocolos de avaliação, de mais do que um codificador das entrevistas leva a maioria dos autores a considerar que a entrevista de revelação deverá sempre ser gravada em vídeo (Wolfe e Birt, 1997; Groz, Kempe e Kelly, 2000; Ney, 1995), ainda que alguns peritos (e.g., Lanning, 2002) considerem que as vantagens e desvantagens deste procedimento devem ser ponderadas. Lanning salienta, entre outras desvantagens, a artificialidade do contexto criado, a frequentemente insuficiente protecção da identidade das vítimas e da confidencialidade das suas declarações e o facto de a atenção, por exemplo em contexto de julgamento, poder ser centrada nos detalhes do modo como a entrevista foi conduzida e não no conteúdo das declarações da criança. Esta é, contudo, uma posição algo minoritária entre os investigadores, havendo mesmo estudos empíricos que mostram a superioridade - do ponto de vista da fidelidade ao que foi dito - da gravação sobre as notas do avaliador (Lamb, Orbach, Sternberg, Hershkowitz e Horowitz, 2000 cit. Bow et al., 2002). Destacamos, ainda, que algumas orientações legais, por exemplo, britânicas (e.g., Criminal Justice Act, 1991) estabelecem que a gravação em vídeo pode substituir o depoimento do menor no Tribunal, conduzindo à recomendação de que esta seja uma prática comum adoptada pelos avaliadores (cf. Memorandum of Good Practice, 1992).

g) Instrumentos de avaliação psicológica

Com já referimos anteriormente (Machado, 2005), um dos dilemas experienciados pelos psicólogos forenses reside na opção entre usar instrumentos

validados mas pouco específicos (e.g., testes ou questionários da área educacional ou clínica) ou recorrer a instrumentos mais adaptados ao contexto da avaliação, mas que carecem de validação adequada. A distinção estabelecida por Otto e Heilbrun (2002) entre "instrumentos de avaliação forense", "instrumentos forensicamente relevantes" e "técnicas clínicas de medida e avaliação" (p.9) é fundamental para esta discussão, por salientar que estas últimas, sendo as melhor validadas e mais usualmente usadas, "tipicamente avaliam construtos (...) que estão consideravelmente distantes da questão específica que se coloca ao decisor judicial" (ibidem, p.9).

Esta é, precisamente, a situação que se verifica no plano da avaliação forense da credibilidade. Apesar do consenso usualmente referido sobre a especificidade técnica exigida para a avaliação do abuso, é frequente o recurso nas avaliações feitas no terreno a instrumentos de avaliação psicológica de largo espectro, tais como o Rorschach, CAT ou WISC. Esta tendência pode, a nosso ver, dever-se a três motivos: a) a carência de instrumentos específicos para a avaliação forense, especialmente se pretendermos que estes estejam validados para a população portuguesa; b) o treino académico recebido pela maioria dos psicólogos no contexto universitário, consideravelmente tradicional, assente nestes instrumentos de largo espectro e ainda insuficientemente sensível aos contextos específicos em que a avaliação é conduzida na prática e, finalmente; c) a convicção, partilhada quer por psicólogos, quer pelos outros operadores do sistema judicial, que o recurso a testes é por si mesmo garantia de uma especial cientificidade no processo de avaliação e que no seu uso reside o cerne do saber psicológico, da sua especificidade e do seu estatuto profissional. Como bem salienta Brodzinsky (1993), "é fundamental que os psicólogos se protejam desta falsa asserção", para a qual "simplesmente não há qualquer evidência empírica" (p.218).

Na verdade, ainda que este tipo de instrumentos - se para eles dispusermos de normas interpretativas adaptadas à população portuguesa - possam revestir algum interesse na compreensão global do caso, é importante salientar que nenhum deles contém qualquer pergunta ou fornece qualquer indicador directo sobre o abuso. Assim, o procedimento usual dos técnicos consiste na comparação (muitas vezes pouco sistemática) do perfil de personalidade / sintomatológico apresentado pela criança com o quadro tipicamente encontrado nas vítimas de abuso, inferindo, a partir da compatibilidade do mesmo, a plausibilidade de aquela experiência ter ocorrido. Algo semelhante, ainda que de forma menos inferencial, é feito com o recurso a instrumentos mais específicos, como sejam as *checklists* de sintomas ou de condutas sexualizadas,

tais como a *Child Sexual Behavior Inventory* (Freidrich et al., 1992 cit. Wolfe e Birt, 1997).

Acontece, contudo, que, como já salientámos anteriormente, "não existe um síndrome clínico específico das crianças abusadas e estas podem apresentar uma grande variedade de sintomas ou mesmo nenhuns" (Machado, 2002, p.50). Desta forma, grande cautela deve ser exercida na utilização da sintomatologia apresentada pela criança para efeitos de ponderação da veracidade. Como afirmam Yuille, Tymofievich e Marxsen (1995), "os indicadores comportamentais de stress devem ser usados apenas como uma parte da recolha de dados sobre um caso e considerados a par dos restantes factos para formular uma conclusão" (p.25). Na mesma linha, Walker (1990) salienta que a confirmação ou desconfirmação das alegações de abuso nunca pode ser provada pela presença ou ausência de quaisquer sintomas.

Conclusões

A análise dos consensos e controvérsias existentes no campo da avaliação da credibilidade das alegações de abuso ilustra claramente, a nosso ver, os desenvolvimentos recentes no campo da avaliação forense, assim como os desafios que esta enfrenta actualmente. Passos fundamentais têm sido dados no sentido de consolidar procedimentos, identificar directrizes, produzir novas propostas de avaliação e desenvolver investigação séria sobre a validade dos métodos usados (Machado, 2005). Ao mesmo tempo, a escassez de instrumentos, as polémicas internas atrás documentadas, a falta de formação específica dos técnicos e a pressão social, judicial e mediática exercida sobre os avaliadores, têm levado a actuações pouco competentes e, por vezes até, pouco éticas.

Neste contexto, torna-se fundamental estimular a investigação, estabelecer parâmetros de procedimentos adequados e divulgar a informação existente. O nosso objectivo com este artigo, tal como com outras publicações recentes (Gonçalves e Machado, no prelo), foi precisamente contribuir para estes objectivos.

Tal não significa, contudo, que acreditemos ser possível chegar a qualquer procedimento de avaliação da credibilidade que seja 100% preciso. Tal como a investigação não produziu nenhum modelo de entrevista que possa ser eleito como o padrão a adoptar sempre (Saywitz, Goodman e Lyon, 2002), não é provável que o consigamos fazer em qualquer outro domínio de avaliação. Partilhamos, por isso, a visão de Conte (1995), segundo o qual é fundamental que os avaliadores estejam precavidos contra os perigos da hiper-confiança,

tantas vezes decorrente da mera ignorância. Qualquer procedimento de avaliação envolve, provavelmente, um julgamento das perdas e ganhos a alcançar com o mesmo e os resultados obtidos devem ser ponderados de acordo com as estratégias utilizadas e seus constrangimentos. Este não é um processo tecnicamente fácil nem corresponde à aplicação de um qualquer algoritmo mais ou menos mágico que por vezes se espera do perito forense. Mas é a única forma - do nosso ponto de vista - de garantir a qualidade e rigor técnico das nossas avaliações, assim como de assegurar a utilidade social do papel do perito.

Referências

- APAV (2002). *Manual CORE. Para o atendimento de crianças vítimas e violência sexual*. Lisboa: APAV.
- Arce, R. e Fariña, F. (2005). El sistema de evaluación global (SEG) de la credibilidad del testimonio: Hacia una propuesta integradora. In R. Arce, F. Fariña e M. Novo (Eds.), *Psicología jurídica* (pp. 103-109). Santiago de Compostela: Xunta de Galicia.
- Barker, P. (1990). *Clinical interviews with children and adolescents*. New York: W.W. Norton e Company.
- Bekerian, D. e Dennett, J. (1995). Assessing the truth in children's statements. In T. Ney (Ed.), *True and false allegations of child sexual abuse* (pp. 163-175). New York: Brunner/Mazel.
- Bow, J., Quinnell, F., Zaroff, M. e Assemany, A. (2002). Assessment of sexual abuse allegations in child custody cases. *Professional Psychology: Research and Practice*, 33, 6, 566-575.
- Brodzinsky, D. (1993). On the use and misuse of psychological testing in child custody evaluations. *Professional Psychology: Research and Practice*, 24, 2, 213-219.
- Bruck, M., Ceci, S. e Francouer, E. (2000). Children's use of anatomically detailed dolls to report genital touching in a medical examination: Developmental and gender comparisons. *Journal of Experimental Psychology Applied*, 6, 1, 74-83.
- Bruck, M., Ceci, S., Francouer, E. e Renick, A. (1995). Anatomically detailed dolls do not facilitate preschoolers' reports of a pediatric examination involving genital touching. *Journal of Experimental Psychology Applied*, 1, 2, 95-109.
- Burkhardt, S. e Rotatori, A. (1995). *Treatment e prevention of childhood sexual abuse. A child-generated model*. Washington: Taylor e Francis.

- Conte, J. (1995). Assessment of children who may have been abused: The real world context. In T. Ney (Ed.), *True and false allegations of child sexual abuse* (pp. 290-302). New York: Brunner/Mazel.
- Davies, G. (1991). Research on children's testimony: Implications for interviewing practice. In C. Hollin e K. Howells (Eds.), *Clinical approaches to sex offenders and their victims* (pp.93-115). Chichester: John Wiley e Sons.
- Ehrenberg, M. F. e Elterman, M. F. (1995). Evaluating allegations of sexual abuse in the context of divorce, child custody, and access disputes. In T. Ney (Ed.), *True and false allegations of child sexual abuse* (pp. 209-230). New York: Brunner/Mazel.
- Gonçalves, R. A. e Machado, C. (Coords.) (no prelo). *Psicologia forense*. Coimbra: Quarteto.
- Groz, C., Kempe, R. e Kelly, M. (2000). Extrafamilial sexual abuse: Treatment for child victims and their families. *Child Abuse e Neglect*, 24, 1, 9-23.
- Hewitt, S. (1999). *Assessing allegations of sexual abuse in preschool children*. Thousand Oaks: Sage.
- Horowitz, S., Lamb, M., Esplin, P., Boychuk, T., Krispin, O. e Reiter-Lavery, L. (1997). Reliability of criteria-based content analysis of child witness statements. *Legal and Criminological Psychology*, 2, 11-21.
- Koocher, G., Goodman, G., White, C., Friedrich, W., Sivan, A. e Reynolds, C. (1995). psychological science and the use of anatomically detailed dolls in child sexual-abuse assessments. *Psychological Bulletin*, 118, 2, 199-222.
- Lanning, K. (2002). Criminal investigation of sexual victimization of children. In J. Myers, L. Berliner, J. Briere, C. Hendrix, C. Jenny e T. Reid (Eds.), *The APSAC handbook on child maltreatment* (pp. 329-347). Thousand Oaks: Sage.
- London, K., Bruck, M., Ceci, S. e Shuman, D. (2005). Disclosure of child sexual abuse. What does the research tell us about the ways that children tell? *Psychology, Public Policy, and Law*, 11, 1, 194-226.
- Machado, C. (2002). Abuso sexual de menores. In C. Machado e R. A. Gonçalves (Coords.), *Violência e vítimas de crimes. Vol II: Crianças* (pp. 39-93). Coimbra: Quarteto.
- Machado, C. (2005). Evoluções paradigmáticas na avaliação forense. *Psicologia: Teoria, Investigação e Prática*, 10, 1, 47-63.
- Machado, C. e Antunes, C. (no prelo). Avaliação de vítimas de abuso sexual. In R. A. Gonçalves e C. Machado (Coords.), *Psicologia forense*. Coimbra: Quarteto.
- Memorandum of Good Practice* (1992). London: The Stationary Office.
- Morgan, M. (1995). *How to interview sexual abuse victims including the use of anatomical dolls*. Thousand Oaks: Sage.
- Ney, T. (1995). Assessing allegations in child sexual abuse: An overview. In T. Ney (Ed.), *True and false allegations of child sexual abuse* (pp. 3-20). New York: Brunner/Mazel.

- Otto, R. e Heilbrun, K. (2002). The practice of forensic psychology. *American Psychologist*, 57, 1, 5-18.
- Ruby, C. e Brigham, J. (1997). The usefulness of the criteria-based content analysis technique in distinguishing between truthful and fabricated allegations. *Psychology, Public Policy and the Law*, 3, 4, 705-737.
- Saywitz, K., Goodman, G. e Lyon, T. (2002). Interviewing children in and out of court: Current research and practice implications. In J. Myers, L. Berliner, J. Briere, C. Hendrix, C. Jenny e T. Reid (Eds.), *The APSAC handbook on child maltreatment* (pp. 349-377). Thousand Oaks: Sage.
- Saywitz, K., Mannarino, A., Berliner, L. e Cohen, J. (2000). Treatment for sexually abused children and adolescents. *American Psychologist*, 55, 9, 1040-1049.
- Veltman, M. e Browne, K. (2002). The assessment of drawings from children who have been maltreated: A systematic review. *Child Abuse Review*, 11, 19-37.
- Vizard, E. (1991). Interviewing children suspected of being sexually abused: A review of theory and practice. In C. Hollin e K. Howells (Eds.), *Clinical approaches to sex offenders and their victims* (pp.117-148). Chichester: John Wiley e Sons.
- Vrij, A. (2005). Criteria-based content analysis. A qualitative review of the first 37 studies. *Psychology, Public Policy and the Law*, 11, 1, 3-41.
- Vrij, A., Akehurst, L., Soukara, S. e Bull, R. (2004). Let me inform you how to tell a convincing story: CBCA and reality monitoring scores as a function of age, coaching, and deception. *Canadian Journal of Behavioural Science*, 36, 2, 113-126.
- Walker, L. (1990). Psychological assessment of sexually abused children for legal evaluation and expert witness testimony. *Professional Psychology: Research and Practice*, 21, 5, 344-353.
- Wolfe, V. e Birt, J. (1997). Child sexual abuse. In E. Mash e L. Terdal (Eds), *Assessment of childhood disorders* (pp.569-623). New York: The Guilford Press.
- Yuille, J., Tymofievich, M. e Marxsen, D. (1995). The nature of allegations of child sexual abuse. In T. Ney (Ed.), *True and false allegations of child sexual abuse* (pp. 21-46). New York: Brunner/Mazel.

CREDIBILITY ASSESSMENT OF SEXUAL ABUSE ALLEGATIONS: CONSENSUS AND CONTROVERSIES

Carla Machado

Departamento de Psicologia, Instituto de Educação e Psicologia, Universidade do Minho

Abstract: Assessment of child victims of sexual abuse is becoming a major area of work for forensic psychologists, with an increasing number of questions being made by judges around the issue of the credibility of abuse allegations. This is an area that has generated great controversy, both nationally and internationally, and one in which major technical and ethical problems arise. This paper intends to map the main areas of consensus within the field of child credibility assessment, as well as to identify and to discuss the major controversies in this area of forensic assessment.

KEY-WORDS: *Child sexual abuse, credibility, forensic assessment.*

CONSTRUÇÃO E AVALIAÇÃO DAS QUALIDADES PSICOMÉTRICAS DO QUESTIONÁRIO DE GESTÃO DO CONHECIMENTO (GC)

Leonor Cardoso
A. Duarte Gomes
Teresa Rebelo

Faculdade de Psicologia e de Ciências da Educação, Universidade de Coimbra, Portugal.

Resumo

O presente artigo focaliza o processo de construção e avaliação das qualidades psicométricas do Questionário de Gestão do Conhecimento (GC), instrumento que visa a avaliação organizacional de processos relacionados com a geração e gestão do conhecimento.

Apresentam-se os procedimentos adoptados na sua construção, bem como aqueles que conduziram à avaliação das suas qualidades psicométricas, designadamente as suas validades de conteúdo e de constructo e fiabilidade (consistência interna das suas dimensões). Analisa-se, ainda, a interpretabilidade da estrutura tetradimensional emergente da análise factorial exploratória realizada no quadro do modelo conceptual adoptado aquando da construção deste instrumento de medida

PALAVRAS-CHAVE: *Avaliação organizacional, construção de instrumentos de medida, avaliação de qualidades psicométricas, gestão do conhecimento, processos organizacionais.*

Introdução

No presente artigo damos conta do processo de construção e avaliação das qualidades psicométricas do GC (acrónimo de *Questionário de Gestão do Conhecimento*) instrumento que tem vindo a ser utilizado em diversas investigações (cf. Brito, 2003; Brito, Gomes e Cardoso, 2005; Cardoso, 2003; Cardoso, 2005), no âmbito das quais se torna necessária a avaliação organizacional de um conjunto de processos relacionado com o conhecimento, com a sua geração e gestão.

Neste contexto, explicitaremos o processo de construção do referido instrumento de medida, apresentando os estudos realizados para avaliar as suas

Morada (address): Leonor Cardoso, Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra, Rua do Colégio Novo, 3001-802 Coimbra, Portugal. Membros do NEFOG – Núcleo de Estudo e Formação em Organização e Gestão (nefog@fpce.uc.pt). A correspondência relativa a este artigo pode ser endereçada para lcardoso@fpce.uc.pt.

qualidades psicométricas, procedimentos através dos quais se procurou garantir a sua posterior utilização em contexto organizacional com um adequado grau de confiança. O GC foi analisado em termos da sua validade (de conteúdo e de constructo) e fiabilidade (consistência interna das suas dimensões). Podemos considerar três tipos de validade: de conteúdo, de constructo e de critério. A validade de conteúdo reporta-se à representatividade e adequação dos itens de um questionário face à variável que pretende medir, bem como à adequação desses mesmos itens à população a que se destinam, em termos de clareza e compreensibilidade. A validade de constructo procura avaliar o grau em que os resultados obtidos com um instrumento são reveladores dos constructos teóricos que lhe estão subjacentes. A validade de critério (não considerada neste trabalho) centra-se na comparação dos resultados obtidos com um determinado instrumento com critérios externos considerados medidas independentes do que o instrumento procura medir (Nunnally, 1978; Fink & Kosecoff, 1985; Hill & Hill, 2000). O objectivo aqui prosseguido tem em vista

(...) oferecer tanto ao “mundo da investigação” como ao “mundo das organizações” instrumentos fiáveis que permitam, no primeiro caso, realizar estudos sobre uma mesma temática, de forma a cumulativamente contribuir para a sua explicação e compreensão e, no segundo, possibilitar um melhor conhecimento de aspectos particulares de uma dada organização, através da sua avaliação, de forma a orientar a tomada de decisão e a intervenção” (Rebelo, Gomes e Cardoso, 2005, p. 191).

Finalizaremos este trabalho analisando a interpretabilidade da estrutura tetradimensional emergente da análise factorial exploratória realizada no contexto do modelo conceptual adoptado aquando da construção deste instrumento de medida.

Construção do GC e sua validade de conteúdo

No processo de construção deste instrumento foram respeitadas as etapas sugeridas por Hill e Hill (2000), designadamente: (a) realização de entrevistas; (b) elaboração de uma versão prévia do questionário; e (c) verificação da sua adequação numa amostra de sujeitos pertencente à população do estudo (pré-teste). A revisão da literatura efectuada foi ainda essencial nesta fase, na medida em que constituiu um elemento orientador fundamental para a análise

do conteúdo do instrumento construído. Neste contexto, sistematizámos e articulámos os contributos teóricos emergentes da literatura revista num modelo que denominámos “Para uma conceptualização e operacionalização da gestão do conhecimento”, no qual sustentámos, do ponto de vista conceptual, a construção dos itens do GC e que pode consultar-se em Cardoso, Gomes e Rebelo (2005).

Etapa 1 (entrevistas)

Foram realizadas seis entrevistas, de carácter exploratório, junto de peritos (no caso concreto, empresários/gestores de topo de organizações industriais do distrito de Viseu), com o objectivo de recolher informação potencialmente relevante para a elaboração dos itens do questionário, nomeadamente ao nível da sua adequação à população a estudar. Estas entrevistas tiveram um carácter semi-estruturado, assente num guião por nós elaborado, que compreendia questões abertas relativas ao tema da gestão do conhecimento. Optámos pela modalidade semi-estruturada da entrevista, pelo facto de esta se adequar aos nossos objectivos, designadamente no que concerne à possibilidade que oferece de serem apresentados os mesmos tópicos aos diferentes entrevistados, permitindo-nos aceder a dados potencialmente comparáveis (Cardoso, 1996, 2000).

As entrevistas foram realizadas após o desenvolvimento de todos os contactos prévios indispensáveis à sua calendarização, tendo cada uma a duração média de uma hora. Após a obtenção da autorização de todos os entrevistados, foram na sua totalidade gravadas em fita magnética, e, posteriormente, integralmente transcritas, tendo sido respeitadas todas as garantias de confidencialidade.

A técnica utilizada para o tratamento dos dados recolhidos foi a análise de conteúdo, tendo sido efectuada uma análise temática baseada nas propostas de Festinger e Katz (1974), Bardin (1979) e Bogdan e Biklen (1994). A aplicação das regras em questão neste caso concreto podem, sinteticamente, expressar-se da seguinte forma: (a) leitura integral de todas as entrevistas com o duplo objectivo de, por um lado, captar o seu sentido global e, por outro, a especificidade de cada entrevista em particular; (b) definição das unidades de análise: unidade de registo (U.R. – foi considerada como unidade de significação a codificar o mais pequeno segmento de texto, pelo que a proposição coincide aqui, com a unidade de registo), unidade de contexto (considerámos cada uma das entrevistas efectuadas, na sua globalidade, como a unidade de contexto que melhor auxiliaria a clarificar a uni-

dade de registo, por exemplo, podermos saber quantos sujeitos referiram determinado aspecto), unidade de enumeração (U.E. – coincide com a unidade de contexto, ou seja, corresponde à entrevista); (c) definição de categorias e subcategorias: o processo de categorização foi efectuado tendo por base as analogias de significado das unidades de registo. Assim sendo, o sistema de categorias é o produto final da progressiva classificação analógica das unidades de registo, pelo que a sua designação conceptual definitiva ocorreu somente no final deste procedimento; (d) teste de validade do sistema de categorização: no sentido de garantir o rigor do processo, submetemos à apreciação de alguns dos elementos da nossa equipa de investigação a análise de conteúdo efectuada, desde o processo de codificação até à definição de categorias.

Os objectivos que nortearam a realização destas entrevistas foram globalmente alcançados, na medida em que nos permitiram, por um lado, aceder à atribuição de sentido efectuada pelos nossos entrevistados ao tema em análise e, por outro, nos facilitaram a tarefa de adequar a linguagem utilizada na construção dos itens do questionário à população a estudar.

Apresentam-se seguidamente os resultados obtidos com a análise efectuada aos conteúdos emergentes das entrevistas individuais realizadas, indicando-se as unidades de registo e de enumeração no contexto das quais foram analisados os diversos indicadores que originaram o sistema de categorias elaborado (cf. Quadro 1).

Quadro 1 – *Gestão do conhecimento*

Categorias		Indicadores	U.R.	U.E.	
Conhecimento do conceito	Sim		3	3	
	Não		2	2	
	Vagamente		1	1	
	Total Categoria		6	6	
Importância atribuída	Muito importante		3	3	
	Importante		1	1	
	Total Categoria		4	4	
Determinantes da sua importância	É importante gerir o conhecimento em prol dos objectivos da organização		4	1	
	Departamentos diferenciados que aprendem coisas diferentes		2	1	
	O conhecimento deve ser posto em prática		1	1	
	O conhecimento organizacional é importante para todos		1	1	
	Estamos muito preocupados com toda a área da informação		1	1	
	Total Categoria		9	4	
Instrumentalidade	Fluidez do conhecimento		1	1	
	Promoção de uma visão mais alargada e sistémica		1	1	
	Melhoria da qualidade do trabalho das pessoas		1	1	
	Total Categoria		3	2	
Sentido atribuído	Gestão do conhecimento das pessoas		7	3	
	Aproveitar o melhor possível os conhecimentos de acordo com as necessidades		2	1	
	Não há possibilidade de gestão económica do conhecimento		2	1	
	Utilização de técnicas para gerir o conhecimento		1	1	
	Gestão das competências geradas a partir do conhecimento		1	1	
	Exercer actividades nas áreas onde se tem melhor conhecimento		1	1	
	Total Categoria		14	4	
Processos	Partilha e disseminação do conhecimento	Reuniões para transmitir informações e aprendizagens		9	3
		Troca de conhecimentos (para, por ex., evitar erros e expandir conhecimentos individuais)		9	1
		Divulgação e disponibilização do conhecimento		6	2
		Abertura para expor ideias		6	1
		Criação de veículos na Intranet		4	1
		Recursos Humanos como veículos privilegiados de transmissão do conhecimento		2	1
		Contactos diários face a face		1	1
		Divulgação do conhecimento por escrito (e.g., jornal da empresa)		1	1
		Transferência via chefias departamentais		1	1
		Subtotal		39	4
	Criação/Aquisição do conhecimento	Formação		10	3
		Identificação e aquisição de conhecimento importante/útil para a organização		5	2
		Participação em colóquios/seminários		2	1
		Melhoria e adaptação dos conhecimentos de base		2	1
		Participação em sessões de apresentação		1	1
		Participação em acções de sensibilização (em pequenos grupos)		1	1
		Divulgação das fontes de informação		1	1
		Leitura de revistas		1	1
	Subtotal		23	4	
Memória organizacional	Conhecimento pode ser armazenado em suportes magnéticos		1	1	
	Subtotal		1	1	
	Total Categoria		63	4	
Total Tema			99	6	

Verificamos que o conceito de “gestão do conhecimento” começa a integrar a prática discursiva destes gestores, uma vez que dos seis sujeitos entrevistados, quatro a ele aludiram (três deles de um modo preciso e um de uma forma mais vaga), sendo a informação detida acerca desta temática decorrente, sobretudo, da leitura de revistas da especialidade. Assim sendo, a análise aqui apresentada focaliza o discurso destes quatro gestores que constituem as quatro unidades de enumeração no âmbito das quais se analisaram as proposições registadas e fizeram emergir o sistema de categorias a este respeito elaborado (com excepção das categorias “instrumentalidade” e “memória organizacional”, onde se registaram apenas três e uma unidades de enumeração, respectivamente).

Constata-se que é grande a importância atribuída à gestão do conhecimento por parte destes gestores, na medida em que a ela se referem como “muito importante”, e “importante”, sendo sobretudo valorizada no âmbito da necessidade sentida de gerir o conhecimento em prol dos objectivos organizacionais. A importância que lhe é atribuída decorre ainda da consciência que os gestores têm de que, por vezes, departamentos e estruturas funcionais distintos aprendem coisas diferentes, importando “concertar” e orientar o aprendido, bem como da imprescindível orientação para a acção de que é portador o conhecimento e que se traduz na preocupação expressa, em termos discursivos, pelos entrevistados, quando referem que “este deve ser posto em prática”.

A análise de conteúdo efectuada possibilitou a categorização da instrumentalidade da gestão do conhecimento (emergente de 3 U.R. e de 3 U.E.), permitindo concluir que, na óptica destes gestores, este processo permite a fluidez do conhecimento por toda a organização, a promoção de uma visão mais abrangente e sistémica por parte de todos os actores organizacionais e a melhoria da qualidade do seu trabalho.

Do discurso dos quatro entrevistados emerge uma atribuição de sentido ao conceito em questão em que este surge associado à gestão do conhecimento detido pelos indivíduos, bem como das competências a partir dele geradas, no sentido de melhor o rentabilizar em função das necessidades organizacionais; é ainda associada à gestão do conhecimento a ideia de que é possibilitadora de um maior investimento organizacional nas áreas onde se detém maior *know-how*. Um dos sujeitos inquiridos enfatizou, neste contexto, a impossibilidade de se gerir economicamente o conhecimento.

O discurso dos entrevistados permitiu identificar práticas ou actividades relacionadas com o conhecimento, no âmbito das quais emergiram aquelas que categorizámos em torno dos processos de partilha e disseminação, cria-

ção/aquisição e memorização. Apresentado em primeiro lugar, porquanto reúne a maioria das unidades de registo a este respeito produzidas (34), o processo de partilha e disseminação do conhecimento surge como aquele que mais preocupa estes gestores ou que mais vulgarmente se associa ao conceito em análise. No caso concreto, operacionaliza-se em termos de um conjunto de actividades, formais e/ou informais, directamente relacionadas com a necessidade de disponibilizar e partilhar informações, conhecimentos e aprendizagens, no sentido de se evitar a prática recorrente de erros já cometidos e de se expandirem as “boas práticas”. No entender destes gestores, é necessário que exista um clima de abertura a novas ideias e, considerando os recursos humanos os veículos por excelência de transmissão do conhecimento, entendem que o seu contacto diário (presencial) se torna fundamental para a concretização do processo em questão. Neste âmbito, é ainda referido o papel essencial das chefias intermédias. Em segundo lugar (reunindo 23 U.R.), surge o processo de criação/aquisição do conhecimento, no contexto do qual a prática discursiva dos gestores se orientou, maioritariamente, para a questão da formação, permitindo concluir que, pelo menos aparentemente, se trata do processo que é percebido pelos sujeitos como principal fonte de criação/aquisição de conhecimento. Consta-se ainda que a valorização do processo em questão (criar/adquirir conhecimento) surge no entanto associada à instrumentalidade (utilidade e importância) do conhecimento a adquirir, sendo esta considerada como critério determinante da sua aquisição; neste âmbito é também referida a necessidade de permanentemente melhorar e adaptar os conhecimentos individuais/organizacionais de base. A divulgação das diversas fontes de informação relevante a todos os actores organizacionais, bem como a sua participação em eventos diversos, tais como sessões de apresentação, de sensibilização, colóquios/seminários e a consulta de revistas da especialidade, são ainda actividades enfatizadas pelos gestores como meios indispensáveis à criação/aquisição do conhecimento. Por último (com 1 U.R. e 1 U.E.), surge a categoria que operacionaliza o processo relacionado com a memória organizacional, percebido como forma de armazenar conhecimento importante e que deve estar disponível para todos os actores organizacionais que dele necessitem. Pode dizer-se que existe aqui uma alusão, ainda que implícita, ao processo de recuperação do conhecimento.

Etapa 2 (versão prévia do GC)

Os itens do GC foram elaborados tendo como referência os resultados a este respeito obtidos na análise de conteúdo das entrevistas realizadas, bem

como os diversos processos relacionados com a gestão do conhecimento (Cardoso, Gomes e Rebelo, 2005), designadamente, (a) criação e aquisição do conhecimento (interna e externa); (b) atribuição de sentido; (c) partilha e disseminação (intencional e não intencional); (d) memória organizacional (interna – intencional e tácita – e externa); (e) recuperação (controlada e automática); e (f) orientação cultural para o conhecimento. Tanto a análise de conteúdo efectuada ao material recolhido nas entrevistas realizadas, quanto o conhecimento prévio que definíamos deste tecido industrial (cf. Cardoso, Rebelo & Brito, 2000; Cardoso, Rebelo & Lourenço, 2001) conduziu-nos à decisão de não considerar na construção desta escala itens relacionados com o processo de medição do conhecimento, dada a sua não aplicabilidade à realidade em questão.

A consideração dos processos acima referidos permitiu-nos igualmente diversificar, tanto quanto possível, o conteúdo dos itens elaborados, em cuja construção e definição das categorias de resposta se procuraram respeitar as indicações de diversos autores (Czaja & Blair, 1996; Fink & Kosecoff, 1985; Ghiglione & Matalon, 1979; Hill & Hill, 2000).

Os itens foram construídos sob a forma de afirmações, tendo sido nossa preocupação que se revestissem de significado para os sujeitos a quem se destinavam, que traduzissem tanto quanto possível o seu contexto profissional quotidiano, que fossem parcimoniosos, isentos de terminologia técnica ou de juízos de valor e que, em termos de construção frásica e de terminologia adoptada, não enviesassem as respostas dos sujeitos. De facto, a necessidade de construir itens de fácil leitura e compreensão (Czaja e Blair, 1996) teve no nosso caso particular relevância, na medida em que sabíamos que uma elevada percentagem dos activos empregados no tecido industrial da região onde seria aplicado o GC era detentora de baixas habilitações escolares. Segundo Cardoso, Rebelo e Brito (2000), mais de metade (64%) dos colaboradores que então integravam os quadros sociais das organizações industriais da região de Viseu possuíam seis ou menos anos de escolaridade. Este facto condicionou igualmente a escolha efectuada por uma escala de tipo Likert com cinco opções de resposta, de acordo com as recomendações de Hill e Hill (2000) para circunstâncias similares às que acabámos de referir.

Tendo em consideração que pretendíamos que os respondentes caracterizassem a organização onde trabalhavam no que concerne aos diferentes processos relacionados com a gestão do conhecimento, optámos por uma escala onde a expressão nuclear é “aplica-se”. Embora conscientes de que se trata de um termo pouco utilizado neste tipo de instrumentos, considerámos que se tratava daquele que mais facilmente conduziria os sujeitos à associação re-

querida das diferentes afirmações ao seu contexto organizacional (*aplica-se nesta empresa, não se aplica*). A versão prévia da escala GC ficou constituída por 56 itens (11 deles invertidos), com cinco opções de resposta: (nesta empresa...) 1 – *nunca se aplica*; 2 – *aplica-se pouco*; 3 – *aplica-se moderadamente*; 4 – *aplica-se muito*; 5 – *aplica-se totalmente*.

Etapas 3 (pré-teste do GC)

Antes de procedermos ao pré-teste deste instrumento, submetemos a versão prévia elaborada à consideração de dois investigadores, detentores de trabalhos realizados em áreas que poderemos considerar como próximas ou afins à da gestão do conhecimento (e.g., aprendizagem organizacional, cultura organizacional orientada para a aprendizagem), de forma a obter a sua apreciação crítica relativamente ao grau de adequação e representatividade dos itens construídos face ao constructo teórico que pretendem medir, respeitando com este procedimento as recomendações de Fink e Kosecoff (1985).

No contexto da análise da validade de conteúdo do GC, encetámos depois a etapa relativa à verificação da adequação e compreensibilidade dos itens construídos, da escala de resposta adoptada, bem como da estrutura global deste questionário à população a estudar (Hill e Hill, 2000), realizando para o efeito o pré-teste numa organização do sector industrial do distrito de Viseu. Esta etapa envolveu duas fases distintas, como sugerem Ghiglione e Matalon (1979): na primeira, o questionário foi aplicado a um grupo de dez pessoas pertencentes a diferentes áreas da organização e com responsabilidades diferenciadas aos níveis das funções desempenhadas; na segunda, aplicou-se o questionário a cinquenta colaboradores da mesma organização, procurando-se, por um lado, avaliar a sua adequação num número mais alargado de pessoas e, por outro, efectuar uma análise simples dos dados recolhidos.

As informações recolhidas na primeira fase conduziram-nos à alteração das opções de resposta relativas aos graus 1 (*nunca se aplica*) e 5 (*aplica-se totalmente*), visto que aproximadamente metade dos respondentes optava por não as seleccionar, considerando-as demasiado restritivas. Procedeu-se por isso à sua alteração, tendo-se optado pela seguinte formulação: *quase nunca se aplica* para o grau 1 e *aplica-se quase totalmente* para o grau 5.

A versão final (já utilizada na segunda fase do pré-teste) da escala de resposta do GC é a seguinte: 1 – *quase nunca se aplica*; 2 – *aplica-se pouco*; 3 – *aplica-se moderadamente*; 4 – *aplica-se muito*; e 5 – *aplica-se quase totalmente*.

No que se refere aos demais aspectos em análise nesta fase de pré-teste, os sujeitos inquiridos consideraram que o questionário era globalmente acessível, tendo apenas aludido a uma certa dificuldade de resposta nos itens invertidos, na medida em que depois de compreenderem o seu sentido, tornava-se difícil estruturar o pensamento de forma a seleccionar correctamente a opção pretendida.

Respeitando uma vez mais as orientações sugeridas por Hill e Hill (2000), o questionário foi, na segunda fase, aplicado em pequenos grupos (de aproximadamente seis pessoas), tendo sido oralmente prestados esclarecimentos relativos ao seu correcto preenchimento e aos objectivos do estudo. No final, os sujeitos verbalizaram um conjunto de opiniões que, regra geral, foram ao encontro do que havia sido referido pelos sujeitos inquiridos na primeira fase.

Procedeu-se depois a uma análise simples dos dados recolhidos em que se avaliaram quais os itens menos respondidos e a distribuição das respostas dadas por item e por questionário. Concluiu-se que os sujeitos respondiam à globalidade dos itens da escala, que as respostas dadas se distribuíam pelas cinco opções de resposta e que, se tivéssemos em consideração a globalidade dos respondentes, as respostas obtidas para um mesmo item não tendiam para a convergência, ou, dito por outras palavras, os sujeitos tendiam a seleccionar diferentes opções de resposta para o mesmo item.

O balanço global dos resultados obtidos com o pré-teste do GC conduziu-nos à decisão de manter a sua estrutura prévia, não tendo portanto sido excluído nenhum dos 56 itens.

Validade de constructo e fiabilidade do GC

As análises da dimensionalidade e da consistência interna do GC aqui apresentadas resultam dos dados relativos a 1275 participantes que integravam os quadros sociais de 50 organizações industriais. De um total de 7020 trabalhadores que integravam os quadros sociais das 50 organizações que constituem a nossa amostra, aquando do processo de recolha de dados, foram distribuídos 1824 questionários e recolhidos 1547. Destes, excluíram-se 272 (17.6%): 267 questionários (17.3%) porque foram devolvidos em branco e 5 questionários (0.3%) porque se encontravam incorrectamente preenchidos. Foram, assim, analisadas as respostas de 1275 sujeitos que apresentavam uma percentagem média de *não-respostas* de 2.89%. Estão, portanto, mais do que garantidas as condições mínimas exigidas para se utilizar a análise factorial, que, de acordo com Bryman e Cramer (1993), Gorsuch (1983) e Pestana e Gageiro (2000), exigem um mínimo de 100 sujeitos por análise e o rácio de

5 para 1, em termos do número de sujeitos por item (no nosso caso, é de 22.8/1). Segundo Bryman e Cramer (1993), o grau de precisão dos factores que emergem de uma análise factorial é função da dimensão da amostra, pelo que se torna essencial que esta seja suficientemente representativa da população de onde foi retirada.

Contudo, antes de se estudar a dimensionalidade deste instrumento, tornou-se necessário efectuar análises preliminares aos dados recolhidos, tendo-se neste contexto efectuado a análise das não-respostas (*missing values*) à totalidade dos questionários recolhidos. Tal como anteriormente referimos, partimos de 1547 casos, eliminámos todos aqueles em que se verificava a existência de mais do que 10% de itens não respondidos (Bryman e Cramer, 1993), que no nosso caso correspondiam aos questionários devolvidos em branco, e todos os que se encontravam incorrectamente preenchidos, tendo-se posteriormente optado pelo método EM (*Expectation Maximization*) para substituição dos restantes *missing values* em variáveis não categoriais (Tabachnick e Fidell, 2001), retendo no final 1275 questionários válidos.

Com o objectivo de tornar a nossa escala mais parcimoniosa e fiável, avaliámos ainda a tendência das respostas dadas pelos respondentes em cada item da escala, analisando, item a item, algumas medidas descritivas (média, desvio-padrão, mínimo, máximo e frequências de resposta por opção), procurando avaliar a pertinência da eliminação de alguns deles. Constatou-se que as respostas obtidas nos 56 itens se distribuíam por todas as opções de resposta (1, 2, 3, 4 e 5) e que somente o item 20 (item invertido) reunia uma percentagem superior a 50% (concretamente, 52%) das respostas dadas pelos sujeitos numa única opção (designadamente, a opção 3 – *aplica-se moderadamente*). Assim, excluindo-se este item, pode-se concluir, de acordo com Matos, Barbosa e Costa (2000), pela existência de uma adequada variabilidade global de respostas por parte dos sujeitos e de um igualmente adequado poder discriminativo dos itens construídos. No que respeita às medidas de tendência central e de dispersão analisadas, verifica-se que, no conjunto dos 56 itens, não existem médias com valores próximos dos extremos da escala, variando os desvios-padrão entre 0.76 e 1.30. Com base nesta análise, de carácter descritivo, apenas o item 20 apresenta características que apontam no sentido da sua exclusão.

Seguidamente, no âmbito da análise da consistência interna do GC, procedemos ao cálculo da correlação entre cada um dos itens e a escala total (através do coeficiente de Pearson), de forma a decidir se através deste procedimento haveria lugar à eliminação de alguns deles por evidenciarem uma correlação inferior a 0,30 com a escala total. Bryman e Cramer (1993) sugere-

rem os critérios propostos por Cohen e Holliday (1982) para a avaliação dos coeficientes de correlação, segundo os quais se considera muito baixa uma correlação inferior a 0,19 e baixa entre 0,20 e 0,30. Com efeito, devido às baixas correlações encontradas em 14 das variáveis, decidimo-nos pela sua eliminação, na medida em que contribuíam para a diminuição da consistência interna do todo (cf. Quadro 2). A análise das correlações inter-item (através do coeficiente de Pearson) corrobora a eliminação destes itens, dado que apresentam correlações nulas ou baixas com metade, ou mais de metade dos itens da escala (Hill & Hill, 2000), sendo a proporção de correlações negativas de 27.5%. De facto, o coeficiente *alpha* de Cronbach inicialmente encontrado (.918), aumentava aquando da exclusão dos itens em questão (cf. Quadro 2). Neste sentido, procedemos ao estudo da dimensionalidade do instrumento com base no conjunto dos restantes quarenta e dois itens.

Quadro 2 – Itens do GC a eliminar: correlações item-escala inferiores a .30 e valores do Alpha de Cronbach da escala total sem os respectivos itens

N.º item	Item item-escala	Correlação sem o item	Alpha total
4	Por vezes ocupamos o lugar de outro colega (porque mudou de trabalho, adoeceu, foi de férias, etc.).	.25	.9178
5	<i>Repetimos os erros do passado.</i>	.23	.9180
10	<i>A função de cada um tem pouco a ver com o que sabe.</i>	.07	.9195
15	<i>As tarefas são realizadas individualmente.</i>	.06	.9196
20	<i>Resolvemos sempre os nossos problemas com soluções novas.</i>	-.48	.9226
25	<i>Os colegas mais experientes guardam para si aquilo que sabem.</i>	.25	.9179
30	<i>Só nós, cá dentro, é que temos informações sobre a nossa empresa.</i>	-.20	.9217
35	<i>Já melhorámos tudo o que tínhamos a melhorar na forma como trabalhamos.</i>	.003	.9198
40	<i>Temos poucas coisas em comum (“falamos linguagens diferentes”).</i>	.12	.9189
41	<i>Os colegas que saem da nossa empresa levam com eles uma opinião sobre nós.</i>	.33	.9182
45	<i>Basta-nos o que já sabemos para continuar a fazer bem o nosso trabalho.</i>	.17	.9186
50	<i>Guardamos só para nós o que vamos fazendo ou aprendendo.</i>	.25	.9178
52	<i>Sabemos que lá fora algumas pessoas sabem se a nossa situação é boa ou má.</i>	.20	.9182
55	<i>Sabemos mais do que é preciso para fazer bem o nosso trabalho.</i>	-.01	.9201

Alpha total = .9180

Nota: o itálico assinala os itens invertidos.

Importa salientar que relativamente ao referido conjunto, os indicadores fornecidos pelo teste KMO (*Kaiser-Meyer-Olkin Measure of Sampling Adequacy*) e pelo teste de esfericidade de Bartlett viabilizam a utilização dos dados da análise factorial, na medida em que, no que concerne ao primeiro, o resultado de 0.94 é um muito bom indicador para a prossecução desta análise, revelando uma amostragem bastante adequada e, relativamente ao segun-

do, a matriz de correlação das 42 variáveis distingue-se da matriz de identidade, indicando intercorrelações, tomadas no seu conjunto, significativamente diferentes de zero [$\chi^2(496) = 1604.81$, $p < .001$]. Segundo Pestana e Gageiro (2000), este indicador, que varia entre 0 e 1, compara as correlações simples com as parciais observadas entre as variáveis. Valores de KMO perto de 1 indicam coeficientes de correlação parciais pequenos, enquanto um coeficiente próximo de 0 alerta para a inadequação da análise factorial, dada a existência de uma correlação fraca entre as variáveis. Estes autores apresentam os seguintes critérios para classificar a relação entre o indicador KMO e a prossecução da análise factorial: $<0,5$, inaceitável a utilização da técnica; entre 0,5 e 0,6, má; entre 0,6 e 0,7, razoável; entre 0,7 e 0,8, médio; entre 0,8 e 0,9, bom; e entre 0,9 e 1, muito bom indicador para a realização da análise factorial. De acordo com aqueles autores, o teste de esfericidade de Bartlett é um outro indicador da existência, ou não, de correlação entre as variáveis. Em conjunto com o KMO permite avaliar a adequação da amostra para a utilização da análise factorial.

Por outro lado, a matriz anti-imagem aponta para a não exclusão de variáveis (itens) e possibilita a aplicação do modelo dos componentes principais (*Principal Component Analysis*), dado que os valores na diagonal principal variam entre .852 e .972 e todos os valores fora desta diagonal são pequenos (o valor máximo absoluto é de .102). A matriz anti-imagem é uma medida de adequação amostral de cada variável para o uso da análise factorial, onde pequenos valores na diagonal principal da matriz (*Measures of Sampling Adequacy* – MSA) levam a considerar a eliminação da variável. Quanto maiores forem estes valores, tanto melhor resulta a análise factorial. Os valores fora da diagonal principal representam o simétrico da matriz de correlações e deverão ser pequenos para a utilização do método dos componentes principais (Pestana e Gageiro, 2000).

O estudo relativo ao número de factores a reter foi efectuado tendo por base o conjunto dos 42 itens, tendo sido respeitado o critério de Kaiser e o *Scree Test* de Cattell (1966), bem como a percentagem de variância total explicada (Matos, Barbosa e Costa, 2000). No entanto, para além dos critérios de ordem estatística, foram ainda considerados outros, de natureza qualitativa, designadamente, de interpretabilidade dos factores a reter em função dos diversos processos que suportam a gestão do conhecimento organizacional e dos resultados obtidos com a realização das entrevistas exploratórias.

Posteriormente, procedemos a uma análise factorial em componentes principais, e, depois de analisadas diferentes alternativas, optámos por uma solu-

ção forçada a quatro factores sujeita a uma rotação ortogonal *varimax*. Importa ainda referir que com o objectivo de que as saturações de cada item com o respectivo factor não suscitasse quaisquer dúvidas, considerámos somente as situadas acima de 0.50 (a única excepção respeita ao item 6, cuja saturação de 0.495, não sendo igual a 0.50, é no entanto muito próxima desse valor). Com base neste critério eliminámos dez variáveis (designadamente, os itens 1, 2, 3, 9, 16, 18, 21, 24, 28, e 51), obtendo um conjunto final composto por trinta e dois itens.

Os quatro factores retidos explicam 49.5% da variância total, e após a rotação, o primeiro apresenta um valor próprio de 5.19, explicando 16.23% da variância, o segundo 4.882 e explica 15.26%, o terceiro 2.99 e explica 9.33%, e o quarto um valor próprio de 2.78, explicando 8.70% da variância. O primeiro factor é saturado acima de .50 por dez itens; o segundo, por onze; o terceiro, por seis e o quarto, por cinco itens.

Estimou-se a consistência interna dos quatro factores utilizando para o efeito o coeficiente *alpha* (α) de Cronbach (cf. *Quadro 3*). O primeiro e o segundo factores apresentam valores de *alpha* de .88 e .86, respectivamente, valores considerados como bons indicadores de consistência interna, dado serem superiores a .80 (Hill e Hill, 2000; Nunnally, 1978). O valor relativo ao terceiro factor é de .79, sendo de .76 aquele que respeita ao quarto factor, indicando graus de consistência interna aceitáveis, na medida em que são superiores a .70 (Hill e Hill, op. cit.). No conjunto dos quatro factores retidos a escala apresenta um coeficiente de consistência interna de .93.

Quadro 3 – GC: saturações factoriais e comunalidades (h^2)

Código	Itens	F1	F2	F3	F4	h^2
GC34	Assistimos a seminários/conferências, lemos o que se publica ou contratamos especialistas	0.754	0.069	0.087	0.145	0.601
GC47	Fazemos circular a informação entre nós (e.g., através de relatórios, jornal interno, e-mail	0.734	0.051	0.177	0.057	0.576
GC37	Juntamo-nos em grupo para resolver alguns problemas	0.709	0.210	0.243	0.100	0.616
GC36	Passamos informação uns aos outros em reuniões de trabalho	0.699	0.242	0.190	0.014	0.583
GC56	Frequentamos cursos de formação ou temos formação no posto de trabalho	0.644	0.071	0.167	0.002	0.448
GC39	Temos meios para registar (guardar) o que de mais importante nós sabemos ou aprendemos	0.616	0.289	0.112	0.193	0.513
GC13	Somos encorajados a tomar a iniciativa	0.606	0.255	0.042	0.303	0.526
GC53	Utilizamos as informações guardadas nos nossos meios informáticos	0.568	0.036	0.132	0.156	0.366
GC12	Colaboramos com outras empresas para adquirir mais informação	0.543	0.043	0.077	0.448	0.503
GC14	São recompensados aqueles que partilham o que sabem	0.524	0.122	0.019	0.341	0.406

Código	Itens	F1	F2	F3	F4	h ²
GC33	<i>Procuramos toda a informação que possa melhorar a qualidade do que fazemos</i>	0.305	0.714	0.015	0.093	0.611
GC44	<i>Todos somos responsáveis pelo que devemos saber para trabalhar com qualidade</i>	0.124	0.711	0.170	0.055	0.553
GC29	<i>Cada um de nós tem uma função a cumprir</i>	0.038	0.653	0.122	0.012	0.443
GC17	<i>O que sabemos vê-se na forma como produzimos</i>	0.069	0.567	0.214	0.273	0.447
GC46	<i>Procuramos estar a par das mudanças que vão acontecendo (e.g., certificação da qualidade)</i>	0.442	0.556	0.085	0.059	0.516
GC31	<i>Pensamos na forma como resolvemos problemas no passado (nos nossos sucessos e insucessos)</i>	0.248	0.534	0.196	0.226	0.437
GC11	<i>Os nossos chefes alertam-nos para o que é importante saber</i>	0.270	0.533	0.064	0.254	0.425
GC07	<i>Sabemos o que se espera de cada um de nós e da empresa (o que nos permite funcionar de um modo coordenado)</i>	0.146	0.532	0.067	0.409	0.476
GC38	<i>Procuramos perceber as regras da nossa organização</i>	0.392	0.527	0.157	0.112	0.469
GC27	<i>Procuramos perceber porque é que correu tão bem determinado trabalho</i>	0.190	0.511	0.278	0.148	0.397
GC32	<i>Agimos de acordo com a forma como estamos organizados</i>	0.062	0.501	0.212	0.152	0.323
GC48	<i>Contamos uns aos outros histórias engraçadas que se passaram no nosso trabalho</i>	0.072	0.064	0.782	0.090	0.629
GC49	<i>Falamos uns com os outros sobre assuntos que não compreendemos bem</i>	0.127	0.268	0.667	0.027	0.534
GC43	<i>Falamos da nossa empresa</i>	0.221	0.296	0.660	0.067	0.577
GC26	<i>Conversamos sobre o trabalho quando casualmente nos encontramos (e.g., a tomar café)</i>	0.133	0.011	0.606	0.108	0.396
GC54	<i>Falamos das nossas funções</i>	0.260	0.292	0.558	0.121	0.479
GC42	<i>Perguntamos aos colegas como resolveram problemas parecidos com os nossos</i>	0.246	0.326	0.508	0.109	0.437
GC19	<i>Sabemos que os nossos concorrentes têm informações sobre nós</i>	0.226	0.013	0.183	0.686	0.555
GC08	<i>Sabemos que os nossos clientes têm uma ideia a nosso respeito</i>	0.218	0.222	0.093	0.662	0.544
GC23	<i>Estamos atentos ao que os nossos concorrentes vão fazendo (e.g., adoptamos os melhores "truques")</i>	0.202	0.322	0.050	0.596	0.503
GC22	<i>O que sabemos é uma "arma" fundamental para ultrapassarmos os nossos concorrentes</i>	0.157	0.410	0.115	0.536	0.493
GC06	<i>O que sabemos vê-se naquilo que fazemos melhor do que os nossos concorrentes</i>	0.036	0.456	0.095	0.495	0.463

Alpha total = .9263

Consideramos que a escala construída e os quatro factores retidos são interpretáveis com base no modelo teórico por nós adoptado (Cardoso, 2005). Nesse sentido, tendo como referência os itens marcadores de cada factor (i.e., aqueles que em cada um deles apresentam maiores saturações factoriais), procedemos à sua denominação, procurando que esta traduzisse o seu sentido global. Ao primeiro factor atribuímos a denominação de "práticas de gestão do conhecimento", ao segundo de "orientação cultural para o conhecimento", o terceiro foi denominado de "gestão social e discursiva do

conhecimento", e por último, o quarto assumiu a denominação de "gestão estratégica do conhecimento".

Sob a designação atribuída, o primeiro congrega os esforços organizacionais desenvolvidos em torno de um conjunto de processos formalmente instituídos que focalizam um conhecimento de natureza não exclusivamente, mas maioritariamente, explícita. Em causa está a tomada de consciência do que se conhece e do que é relevante conhecer e a adopção de um conjunto de práticas que visam fundamentalmente a criação/aquisição de novo conhecimento, a sua preservação, partilha e utilização, que ocorre sobretudo pela imprescindível incorporação do conhecimento detido em produtos e serviços. A prática adoptada, as rotinas e os procedimentos instituídos implicam, portanto, a prévia detenção de conhecimento, a consciência dessa mesma detenção e, forçosamente, a sua utilização ou, de outro modo dito, importa que se conheça, que se saiba que se conhece e que se utilize o conhecido. A mudança e o desenvolvimento organizacionais tornam-se possíveis porquanto a consciência do conhecimento detido acerca de algo possibilita a crença e a vontade de se conhecer mais e melhor, viabiliza a tomada de iniciativa, a improvisação e a inovação.

O segundo reflecte um quadro que serve de orientação para as práticas, regras, normas e procedimentos instituídos (focalizados no factor 1). Trata-se de um referencial comum, de uma memória colectiva, onde é guardado e preservado o que se revelou instrumental ou relevante e que se sabe utilizar. Nesta orientação é acentuada a importância acerca de onde vimos, onde estamos e para onde queremos ir; nela é especificado um conjunto de valores, partilhado e considerado basilar, que é conhecido e reconhecido. Neste sentido, pode falar-se de uma cultura, de uma matriz interpretativa, a partir da qual se destriçam ameaças e oportunidades, se perspectivam problemas solúveis e insolúveis, se formula um entendimento quanto ao que é exequível ou impraticável, se delimita o que é ou não relevante. Falamos de um quadro, de um "pano de fundo", inscrito numa memória, que orienta e atribui sentido às práticas que se adoptam no dia a dia e que traça uma certa orientação. Falaremos, neste caso concreto, da gestão de uma dada cultura, mais especificamente de uma "cultura orientada para o conhecimento". Nesta cultura, a qualidade e a competitividade ocupam um lugar central. Trata-se de uma cultura que é conhecida e reconhecida, onde é nuclear o valor do conhecimento no sentido da produtividade, da qualidade e da competitividade.

O terceiro factor traduz as interacções que, ao nível informal, ocorrem na organização e que facilitam a construção social do conhecimento, no âmbito

da qual é determinante a prática discursiva e a criação de uma linguagem comum. Em causa está um conhecimento de natureza maioritariamente tácita e que dificilmente se cria e se põe em prática em circunstâncias diferentes destas. Trata-se de uma gestão simbólica do conhecimento que possibilita a atribuição de sentido ao que se conhece e se faz, ao que os outros conhecem e fazem, em que, falando ao e para trabalhar, se constroem compreensões ou elaboram entendimentos, que se pretendem colectivos, acerca de acontecimentos organizacionais relevantes.

Por fim, o quarto factor reflecte uma orientação para o exterior por parte da organização. Esta orientação focaliza-se na utilização do conhecimento organizacional numa lógica de competitividade e no âmbito da qual os clientes e os concorrentes assumem um papel determinante. Em causa está uma gestão estratégica do conhecimento, onde se encontra permanentemente em jogo o conhecimento que detemos, o que é detido por outros, o que queremos (ou não) que se saiba que sabemos, o que outros pretendem (ou não) que saibamos e que visa, antes de mais, uma resposta organizacional cada vez mais eficaz a todas as suas "partes interessadas", a adaptação ao ambiente externo e a promoção de uma vantagem competitiva caracterizada pela sustentabilidade.

Em jeito de síntese, pode afirmar-se que esta escala e os quatro factores nela retidos são interpretáveis à luz da literatura revista sobre gestão do conhecimento e do modelo por nós proposto (Cardoso, Gomes e Rebelo, 2005). Efectivamente, se algo de novo emerge nas questões relacionadas com o conhecimento tem a ver com o facto de que a sua gestão é, antes de mais, saber, deter conhecimento, ter dele consciência (saber o que se sabe) e utilizá-lo ou pô-lo em prática. Este é o entendimento que orienta e dá sentido a esta escala. Assim sendo, é com base no que conscientemente se conhece que se torna possível adoptar um conjunto de práticas (de gestão do conhecimento) que, em última instância, propicia a criação e utilização de novos saberes, catalisadores da criatividade e da inovação. Contudo, estas práticas pressupõem a existência de uma memória (individual e colectiva) com base na qual uma certa orientação é adoptada e um dado conjunto de valores emerge como nuclear. Uma cultura (de qualidade e orientada para o conhecimento) conhecida, reconhecida e partilhada dá forma e enforma as práticas adoptadas, assim adquirindo sentido e inteligibilidade. Neste contexto, assumem particular importância os processos relacionados com a construção social e discursiva do conhecimento, com a sua gestão simbólica, na medida em que são fundamentais para a atribuição de um sentido colectivo à realidade organizacional. A sobrevivência das organizações, a sua competitividade sus-

tentada e sustentável, principal objectivo prosseguido, é simultaneamente catalisada pela concretização dos objectivos internamente orientados (em jogo em qualquer dos 3 factores anteriores), mas profundamente dependente de uma orientação externa, para o mercado, para os clientes ou destinatários da acção prosseguida, configurada numa gestão do conhecimento de cariz estratégico.

Conclusão

Neste artigo centrámo-nos no processo de construção e avaliação das qualidades psicométricas do GC. Os resultados obtidos nos estudos realizados e aqui apresentados revelaram-se positivos e encorajadores. O objectivo de explicitar as etapas percorridas na construção de um instrumento de medida susceptível de contribuir para o avanço da investigação empírica relacionada com a gestão do conhecimento julgamos ter sido conseguido. Importa, no entanto, salientar que os objectivos associados ao nosso interesse pela gestão do conhecimento não se restringem a este aspecto e que o referido instrumento é passível de outras utilizações, podendo, por exemplo, ser inserido no contexto da avaliação organizacional, ser encarado como um meio que possibilita o diagnóstico ou facilita a intervenção a este nível. Algumas das suas potencialidades e limitações devem ainda ser assinaladas. E, para concluir, consideramos que, pensando no contexto educativo ou escolar, algumas implicações pedagógicas merecem ser explicitadas.

Estamos, pois, conscientes de que estudos adicionais se tornam necessários para que as qualidades agora evidenciadas por este instrumento possam vir a confirmar-se no futuro em similares e distintos contextos organizacionais. Assim sendo, e tendo em consideração o atrás referido, parece-nos relevante estudar o comportamento do instrumento em questão no contexto específico das organizações escolares/educativas. Inserida numa linha de investigação centrada sobre o estudo dos processos organizacionais relacionados com a aprendizagem e o conhecimento, encontra-se actualmente em desenvolvimento uma investigação que visa, precisamente, adaptar e validar o GC para organizações escolares. Trata-se de um aprofundamento do presente trabalho e dos resultados da investigação em que estamos agora empenhados esperamos dar conta em futuras publicações.

Ao concluir, faz sentido mencionar algumas implicações pedagógicas do trabalho já realizado, inserindo-as no contexto do ensino/investigação de processos relacionados com o conhecimento organizacional.

Começaremos por enfatizar o nível de análise/intervenção onde nos posicionamos, o organizacional, por contraste com o individual, nível a partir do qual as questões relacionadas com contextos educativos/pedagógicos são, habitualmente, perspectivadas. De facto, habitualmente, as organizações escolares consideradas como excelentes, aquelas que são bem cotadas ou estão bem posicionadas no dito “ranking”, ou seja, que são avaliadas como eficazes, vêem frequentemente o seu êxito associado à detenção de bons professores e de bons alunos. Em nosso entender, esta é uma visão reducionista da questão, porquanto insuficiente e limitada para a determinação dos critérios/indicadores do bom desempenho das organizações educativas. Para além dos conhecimentos individuais e dos êxitos individuais, as escolas diferenciam-se, afirmam-se e alcançam bons resultados quando estão dotadas de um corpo docente profissionalizado e estabilizado, quando têm uma liderança e um projecto, quando sabem onde estão e aonde querem chegar, quando têm consciência da sua razão de ser (missão), conhecem os destinatários da sua acção e adquirem aos seus olhos a legitimidade de que carecem para sobreviver e ter êxito. O que acaba de ser dito – e que diversos relatórios oficiais sobre o sistema de ensino enfatizam – tem que ver com o que actualmente vem sendo chamado “gestão do conhecimento” e de que aqui nos ocupámos. Uma organização que sabe que sabe, que tem consciência (conhecimento) do que sabe e o valoriza, que conhece quem no seu seio sabe o quê, onde está, onde o procurar, como o obter e como o partilhar, dispõe de um “trunfo”, ou de um “capital intelectual”, que é importante para assegurar o seu futuro e ter êxito, para enfrentar as ameaças e aproveitar as oportunidades. O “conhecimento organizacional” que gera e o modo como o sabe gerir em muito contribui para que os seus resultados sejam o que são. Diferentes tipos ou formatos de organização permitem que os indivíduos que as habitam e as viabilizam tenham diferentes tipos de pensamento e obtenham resultados distintos. As organizações não se constroem a si mesmas nem são exteriores às pessoas que as fazem e as tornam eficazes. Contudo, à medida que construímos as organizações em que vivemos, e sem as quais não viveríamos como vivemos, vamos moldando a forma como pensamos e agimos e dotando-as da legitimidade de que elas carecem para sobreviver. E as organizações que consideramos legítimas tomam decisões que nos afectam e condicionam. Saber distinguir o que uma organização faz por rotina e sem pensar (conhecer a cultura que a identifica) e o que faz sabendo que sabe o que está a fazer e tirando partido do que sabe (conhecimento organizacional) poderá constituir uma vantagem competitiva e, por isso, vale a pena prestar-lhe a atenção devida e ocupar-se da gestão desse conhecimento.

As organizações, sejam elas escolares ou empresariais, ganham em caminhar no sentido da definição de estratégias que as levem, cada vez mais, a “incorporar” não só o seu ambiente interno, mas, igualmente, a sua envolvente externa na definição das suas estratégias de mudança e inovação. A decorrente “orientação para o exterior”, emergente deste novo olhar para a gestão das organizações escolares, pressupõe a capacidade de utilização de um conhecimento – de natureza organizacional e não apenas individual – que é susceptível de rentabilizar o conhecimento detido, bem como de facilitar a aquisição/criação daquele que se afigura como instrumental, mas que ainda se não detém. Gerir estrategicamente as organizações educativas, passa, assim, por adquirir e utilizar competências de gestão estratégica de um conhecimento de natureza organizacional.

Referências

- Bardin, L. (1979). *Análise de conteúdo*. Lisboa: Edições Setenta.
- Bogdan, R., e Biklen, S. (1994). *Investigação qualitativa em educação: Uma introdução à teoria e aos métodos*. Porto: Porto Editora.
- Brito, E. J. O. (2003). Gestão do conhecimento e qualidade: Estudo empírico sobre a gestão do conhecimento em câmaras municipais que adoptaram a gestão pela qualidade. Tese de Mestrado não publicada. Faculdade de Psicologia e de Ciências da Educação, Universidade de Coimbra.
- Brito, E., Gomes, A. D., e Cardoso, L. (2005). Gestão do Conhecimento: Adaptação e validação da escala GC para o sector autárquico. *Psychologica*, 38, 9-22.
- Bryman, A., e Cramer, D. (1993). *Análise dos dados em ciências sociais: Introdução às técnicas utilizando o SPSS*. Oeiras: Celta. (Obra original publicada em 1990).
- Cardoso, L. (1996) *Aprendizagem organizacional: Contributo para uma metodologia de diagnóstico: Perfil de Aprendizagem*. Tese de Mestrado não publicada. Faculdade de Psicologia e de Ciências da Educação, Universidade de Coimbra.
- Cardoso, L. (2000). Aprendizagem organizacional. *Psychologica*, 23, 95-117.
- Cardoso, L. (2003). Gerir conhecimento e gerar competitividade: Estudo empírico sobre a gestão do conhecimento e seu impacto no desempenho organizacional. Dissertação de Doutoramento em Psicologia. Especialização em Psicologia do Trabalho e das Organizações, apresentada à Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra. Disponível no site da Ariadne Editora. (www.ariadne-editora.com).

- Cardoso, L., Gomes, A. D., e Rebelo, T. (2005). Para uma conceptualização e operacionalização da gestão do conhecimento. *Psychologica*, 38, 23 – 44.
- Cardoso, L., Rebelo, T., e Brito, E. (2000). *Recursos humanos do sector industrial do distrito de Viseu*. Viseu: AIRV – Associação Empresarial da Região de Viseu.
- Cardoso, L., Rebelo, T., e Lourenço, F. (2001). Organizações industriais em aprendizagem: A região de Viseu em perspectiva. Viseu: AIRV – Associação Empresarial da Região de Viseu.
- Cardoso, P. (2005). Gestão do conhecimento em organizações hoteleiras do distrito de Viseu. Tese de Mestrado não publicada. Departamento de Economia, Gestão e Engenharia Industrial. Universidade de Aveiro.
- Cohen, L., e Holliday, M. (1982). *Statistics for social scientists*. London: Harper e Row.
- Czaja, R., e Blair, J. (1996). *Designing surveys: A guide to decisions and procedures*. Thousand Oaks, CA: Pine Forge Press.
- Festinger, L., e Katz, D. (1974). *Les méthodes de recherche dans les sciences sociales*. Paris: Presses Universitaires de France.
- Fink, A., e Kosecoff, J. (1985). *How to conduct surveys: A step-by-step guide*. Newbury Park: Sage Publications.
- Ghiglione, R. e Matalon, B. (1979). *Les enquêtes sociologiques: Théories et pratique*. Paris: Armand Colin.
- Gorsuch, R. (1983). *Factor analysis*. Hillsdale, NJ: Lawrence Erlbaum.
- Hill, M., e Hill, A. (2000). *Investigação por questionário*. Lisboa: Edições Sílabo.
- Matos, P., Barbosa, S., e Costa, M. (2000). A propósito da avaliação psicológica da identidade: Estudo sobre o EOMEIS-2. *Psychologica*, 25, 5-17.
- Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill.
- Pestana, M. H., e Gageiro, J. (2000). *Análise de dados para ciências sociais: A complementaridade do S.P.S.S.* (2ª ed.). Lisboa: Edições Sílabo.
- Rebelo, T., Gomes, A. D., e Cardoso, L. (2005). Cultura de aprendizagem: A (bi)dimensionalidade da escala OCA. *Psychologica*, 38, 191-208.
- Tabachnick, B. G., e Fidell, L. S. (2001). *Using multivariate statistics* (4ª ed.). Boston: Allyn and Bacon.

CONSTRUCTION AND EVALUATION OF PSYCHOMETRIC QUALITIES OF KNOWLEDGE MANAGEMENT QUESTIONNAIRE

Leonor Cardoso
A. Duarte Gomes
Teresa Rebelo

Faculdade de Psicologia e de Ciências da Educação, Universidade de Coimbra, Portugal.

Abstract: The present article focuses on the process of construction and evaluation of the psychometric qualities of Knowledge Management Questionnaire, an instrument aiming towards an organizational assessment of processes concerned with the generation and management of knowledge. The procedures adopted in its construction are presented, as well as those that led to the evaluation of its psychometric qualities, namely its content and construct validity and reliability (internal consistency of its dimensions). We then analysed the interpretability of the tetra-dimensional structure, emergent from the factorial exploratory analysis accomplished in the conceptual model adopted at the time of the construction this instrument of measure.

KEY-WORDS: *Organizational assessment, construction of instrument of measure, evaluation of psychometric qualities, knowledge management, organizational processes.*

Barros de Oliveira, J. H. (2005). *Psicologia da Educação* (1º e 2º vol.). Porto: LivPsic.

Em 1996, o autor, que há muito rege a cadeira de Psicologia da Educação na Faculdade de Psicologia e de Ciências da Educação da Universidade do Porto, publicou, juntamente com A. Barros, uma espécie de Manual de apoio às aulas, intitulado *Psicologia da Educação Escolar*, em dois volumes, o 1º dirigido mais à Aprendizagem (Aluno) e o 2º ao Ensino (Professor). Os dois volumes esgotaram-se e saiu 2ª edição em 1999 que, entretanto, também se esgotou. José Barros - agora único autor, uma vez que, entretanto, A. Barros faleceu - resolveu reeditá-lo também em dois volumes, mantendo os mesmos capítulos mas actualizando muitos temas e a bibliografia que figura abundantemente ao fim de cada volume. O livro intitula-se agora simplesmente "Psicologia da Educação", título há muito consagrado.

Eis os capítulos do 1º volume: 1. Conceito, estatuto e história da psicologia da educação; 2. Inteligência, memória e aprendizagem; 3. Motivação, personalidade e aprendizagem; 4. Dificuldades de aprendizagem – dislexia – deficiência mental – educação especial; 5. Insucesso escolar; 6. Sobredotação e criatividade. No 2º volume trata os temas seguintes: 1. Personalidade, expectativas e eficácia do professor; 2. Dinâmica de grupo e disciplina na sala de aula; 3. Relação educativa; 4. Desenvolvimento psicológico e educação; 5. Objectivos educativos, avaliação e observação; 6. Psicologia escolar (competências do psicólogo na escola). Ninguém duvida do interesse e pertinência destes temas.

Enquanto o 1º volume se dirige essencialmente ao aluno e à sua aprendizagem, o 2º centra-se no professor e no ensino. Mas é evidente que se trata de dois pólos educativos inseparáveis e interagentes. Aliás, como afirma o autor, a Psicologia da Educação tem como núcleo central o processo dinâmico ensino-aprendizagem. Este processo pode vir a ser mais enriquecido se os educadores, e não apenas os psicólogos, dominarem algumas estratégias de ensino-aprendizagem e tiverem algum (ou muito, como é desejável) conhecimento da pessoa do educando, cientes de que não se dirigem apenas à 'cabeça' dos alunos, aos aspectos cognitivos, mas também ao 'coração' ou à sua emoção e motivação, preocupando-se com a educação integral.

O autor termina a breve apresentação da obra, afirmando: "Oxalá este livro possa dar o seu modesto contributo para um ensino-aprendizagem mais eficientes e uma educação que faça mais felizes educadores e educandos, contribuindo assim para um mundo melhor, no dealbar do século XXI e do terceiro milénio". Ao mesmo tempo que fazemos nossos estes votos, recomenda-

mos este Manual principalmente a estudantes de Psicologia e de Ciências da Educação, aos psicólogos educacionais e aos professores de qualquer nível de ensino.

Félix Neto

Barros de Oliveira, J. H. (2005). *Psicologia do envelhecimento e do idoso* (2ª ed.). Porto: LivPsic.

Em menos de meio ano esgotou-se a 1ª edição deste livro, o que diz do seu interesse e aceitação junto do público que se interessa pelos temas aí versados e que reflectem uma problemática muito actual. No mesmo ano da 1ª edição, sai a 2ª, não vendo o autor necessidade de fazer alterações em tão curto espaço de tempo.

O livro tenta, desde a perspectiva psicológica, abranger diversas abordagens do envelhecimento, privilegiando a desenvolvimental. Este Manual – podemos designá-lo assim – sobre a psicologia do envelhecimento e do idoso distribui-se ao longo de oito capítulos: questões preliminares, interpretação do envelhecimento, perspectiva desenvolvimental do envelhecimento, outras perspectivas personológicas do envelhecimento, emoções positivas e valores do idoso, saúde e doença mental na senilidade, o idoso na família e na sociedade, o idoso e a morte. No final, uma abundante e actual bibliografia ajuda a aprofundar qualquer um dos temas versados.

O autor afirma, na Introdução, que “se sentiria compensado se o livro proporcionasse um melhor conhecimento do processo de envelhecimento e da psicologia do idoso, facilitando aos mais novos uma maior compreensão e respeito pelos gerontes e a estes uma maior qualidade de vida” (p. 8). São estes também os nossos votos.

Félix Neto

Fotocomposição e impressão:

Claret - Companhia Gráfica do Norte – Rua do Colégio, 69

4415-284 Pedroso – Tel. 22 787 73 20 - Fax 22 787 73 29

INSTRUÇÕES AOS AUTORES

1. A revista aceita artigos originais no domínio da investigação psicológica, da educação, da cultura e das práticas educativas inovadoras. Os autores deverão expressamente declarar na carta ao Director que o artigo é original e que não foi objecto de qualquer publicação anterior. Os artigos serão objecto de avaliação por especialistas doutorados nas áreas respectivas.

2. A revista publica dois números por ano, o 1º em Maio e o 2º em Dezembro. O prazo limite para publicação no número de Maio é 31 de Janeiro e no número de Dezembro é 31 de Julho. Nas primeiras quinzenas de Março ou de Outubro os autores serão informados da aceitação ou não do artigo, das correcções a introduzir e do envio de uma disquete com a versão final corrigida.

3. Os artigos a submeter devem ser enviados em triplicado em cópias laser e por norma não devem exceder 20 folhas A4 redigidas a 2 espaços. A 1ª folha deve conter o título, o nome dos autores, a instituição e o endereço para correspondência. A 2ª folha deve incluir o resumo e as palavras-chave em português. A 3ª folha e seguintes incluirá o corpo do artigo que deverá concluir com uma listagem ordenada das referências bibliográficas citadas. Na folha a seguir às referências seguem-se por ordem as notas, os quadros, as figuras e diagramas. A última folha incluirá a versão em inglês do título, instituição, resumo e palavras-chave para efeitos de indexação em bases de dados internacionais. O resumo em português e em inglês não deve exceder as 120 palavras. Os autores devem evitar o «bold» e os sublinhados no texto e reduzir ao mínimo as notas de pé-de-página.

4. Os títulos e secções do artigo não devem ser precedidos por números, têm maiúsculas na 1ª letra das palavras e seguem o formato seguinte: 1ª ordem: Tipo normal, centrado; 2ª ordem: Tipo normal, indexado à esquerda; 3ª ordem: Tipo normal, indexado ao parágrafo; 4ª ordem: Tipo itálico, indexado ao parágrafo.

5. As referências bibliográficas devem ser elaboradas de acordo com as normas de "Publication Manual of APA (1994, 4ª ed)" com algumas adaptações para português, nomeadamente a substituição do "&" por "e", "(2nd ed.) por

(2ª ed.)", "(3rd vol.) por (3ª vol.)" conforme a nota seguinte.

6. As referências mais frequentemente usadas para artigo, livro, capítulo, livro traduzido e comunicação oral (paper) devem obedecer ao formato dos exemplos seguintes:

Artigo: Recht, D. R., e Leslie, L. (1988). Effect of prior knowledge on good and poor readers' memory of text. *Journal of Educational Psychology*, 80, 16-20.

Livro: Skinner, B. F. (1971). *Beyond freedom and dignity*. New York: Knoff.

Capítulo em livro: Neisser, U., e Harsch, N. (1992). Phantom flashbulbs: False recollections of hearing the news about Challenger. In E. Winograd e U. Neisser (Eds.), *Affect and accuracy in recall: Studies of "Flashbulb memories"* (pp. 9-31). Cambridge: Cambridge University Press.

Livro traduzido: Skinner, B. F. (1974). *Para além da liberdade e da dignidade* (J. L. D. Peixoto, trad.). Lisboa: Edições 70. (Obra original publicada em 1971). No corpo do artigo deve referir-se, Skinner (1971/1974).

Comunicação oral: Taylor, M. (1996, Agosto). *Post-traumatic stress disorder, litigation and the hero complex*. Comunicação oral apresentada no XXVI Congresso Internacional de Psicologia, Montréal, Canadá.

7. Quando no corpo do artigo são citados autores, cuja investigação foi conhecida indirectamente através de outros autores, deve proceder-se assim: No corpo do artigo escreve-se: «Segundo Godden e Baddeley, citado por Zechmeister e Nyberg (1982, p. 123), ... »; Nas referências cita-se apenas o autor que foi lido directamente, Zechmeister e Nyberg (1982).

8. Os Quadros e as Figuras devem ser sequencialmente ordenados em numeração árabe ao longo do texto. A legenda do Quadro deve estar escrita por cima e a da Figura ou Diagrama por baixo.

9. Os artigos são da inteira responsabilidade dos seus autores. Os artigos aceites para publicação ficam propriedade editorial da Revista. Qualquer reprodução integral ou parcial dos mesmos apenas pode ser efectuada após autorização escrita do Director.

10. Os autores recebem 3 exemplares da revista em que um ou mais trabalhos seus sejam publicados. Não serão feitas separatas dos artigos.

LIVROS RECEBIDOS E RECENSÕES

A Revista fará uma listagem dos livros enviados pelas Editoras no segundo número de cada ano. Os autores ou editores, que desejarem a publicação de recensões, deverão enviar dois exempla-

res da obra em causa. O Conselho Editorial reserva-se o direito de publicar apenas as recensões das obras que se enquadrem nos objectivos da Revista.

