

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*

Mário R. Simões**

Andreia Azevedo***

Belmira Marques***

Catarina Morgado***

Conceição Costa***

Luís Marques***

Ana Filipa Lopes****

Resumo: As versões portuguesas da WISC-III e da WISC foram administradas, de forma contra-balanceada no que respeita à ordem de apresentação, a uma amostra de 70 crianças e adolescentes, observados em contexto clínico e com idades compreendidas entre os 6 e os 15 anos (mediana=10 anos; 67% do sexo masculino e 33% do sexo feminino). Do ponto de vista clínico, este grupo é bastante heterogéneo. As correlações elevadas entre os QIs Verbal (.77), de Realização (.74) e da Escala Completa (.78) de ambas as escalas constituem uma confirmação de que a WISC-III mede essencialmente os mesmos constructos que a WISC.

As pontuações QIEC obtidas na WISC-III são, em média, aproximadamente, 28 pontos inferiores às alcançadas na WISC. Os resultados no QIV e no QIR na WISC-III estão situados, respectivamente, 18 e 31 pontos abaixo dos resultados no QIV e no QIR da WISC.

Palavras-chave: Avaliação psicológica; Inteligência; Validade; WISC-III; WISC.

Abstract: The WISC-III (2003) and WISC (1969) portuguese standardisations were administered in counterbalanced order to a mixed clinical sample (mainly with learning disabilities and/or with behavioural and emotional difficulties) of 70 children aged 6–15 years (median = 10 years). The sample included 23% learning disabilities diagnosis, 70% comorbidity conditions (learning disabilities, opposition, attention deficit hyperactivity disorder, anxiety or depression), 3% family problems, 1% epilepsy and 3% with ill-defined problems.

* Este trabalho insere-se no contexto de um projecto de investigação financiado pela Fundação para a Ciência e a Tecnologia (FCT) cujo objectivo é a “Aferição nacional das escalas de Inteligência de Wechsler (WPPSI-R e WISC-III) para a população portuguesa” (PRAXIS/PCSH/PSI/C/91/96) e que contou com o apoio da CEGOC. A redacção deste texto beneficiou igualmente do apoio do projecto de investigação “Aferição de testes neuropsicológicos para a população portuguesa” (SAPIENS/POCTI/PSI/35410/2000) e do Centro de Psicopedagogia da Universidade de Coimbra (Unidade financiada pela FCT, Linha de Acção n.º 2: “Adaptação e aferição de testes psicológicos e pedagógicos para a população portuguesa”). Os autores agradecem a disponibilidade e o acolhimento do Director do PSIANIMA, Dr. Rui Ramos.

** Serviço de Avaliação Psicológica. Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra. Rua do Colégio Novo, 3001-802 Coimbra. [e-mail: simoesmr@fpce.uc.pt]

*** Psicólogas/Psicólogo. PSIANIMA, Centro de Estimulação e Consulta Psicológica. Espinho.

**** Bolseira do Projecto SAPIENS/POCTI/PSI/35410/2000 (FCT).

The correlations between Verbal Intelligence Quotient, Performance Intelligence Quotient and Full Scale Intelligence Quotient are substantial (.74 to .78). The magnitude of these correlations provides some evidence that WISC-III measures essentially the same constructs as does the WISC.

A comparison of the mean scores shows that WISC-III FSIQ is approximately 28 points less than the WISC FSIQ score and that WISC-III VIQ and PIQ scores are approximately 18 and 31 points less than the corresponding WISC scores respectively. These differences between WISC-III and WISC scores are expected, due to the number of years between the standardisation of the two tests (see Flynn effect), but more pronounced to those found between international data concerning successive revisions of the WISC scales (WISC-III and WISC-R scales). WISC-R scale was never adapted for Portuguese population.

More research is needed. Particular emphasis should be placed on using normal children and adolescents samples in order to see if a similar magnitude of results is obtained for a non-clinical sample.

Key words: psychological assessment; intelligence; validity; WISC-III; WISC.

Introdução

As Escalas de Inteligência são instrumentos importantes na avaliação e diagnóstico em contextos educativos e clínicos, sendo frequentemente usadas em situações de reavaliação, monitorização e medida da eficácia das intervenções. Em contexto educativo, os resultados nas Escalas de Inteligência, incluindo os QIs, constituem elementos importantes na avaliação e diagnóstico de casos de Deficiência Mental, Dificuldades de Aprendizagem e Sobredotação.

As Escalas de Inteligência de Wechsler para Crianças (WISC, WISC-R e WISC-III) encontram-se entre as provas de avaliação mais usadas em todo o mundo. A WISC-III, originalmente publicada em 1991, nos Estados- Unidos, conta com aferições em numerosos países como é o caso, a título de exemplo, da Inglaterra (1992), França/ Bélgica (1996) ou Brasil (2003). A aferição portuguesa da WISC-III acaba igualmente de ser disponibilizada (Wechsler, 2003) e irá substituir a WISC. De um modo geral,

a WISC tem sido a versão habitualmente usada pelos psicólogos portugueses na avaliação psicológica da criança e do adolescente com idades compreendidas entre os 6 e os 15 anos.

Uma das questões que atravessa o processo de aferição de uma escala de inteligência, neste caso da WISC-III, remete justamente para o domínio da validade dos seus resultados e, neste particular, para exercícios de comparação com os resultados noutras escalas de inteligência. A confrontação entre resultados obtidos nas Escalas de Inteligência de Wechsler pode assumir formas e objectivos diversos: (i) comparação entre as diferentes versões de uma mesma escala (p. ex., WISC, WISC-R, WISC-III) em estudos de validade concorrente, quando se pretende fazer a transposição da WISC para a WISC-III, conhecer o significado “actual” dos resultados obtidos com uma versão mais antiga, como é o caso da WISC; (ii) cotejo entre versões distintas de escalas diversas, mas com algumas idades de sobreposição (p. ex., entre os 6 anos e os 6 anos e 6

meses, no caso da WPPSI-R e da WISC-III; ou na idade de 16 anos, na WISC-III e WAIS-III), igualmente em estudos de validade concorrente; ou (iii) comparação entre os resultados obtidos por um mesmo sujeito, em momentos distintos do seu desenvolvimento, quando ele é examinado sucessivamente a partir de diferentes escalas (p. ex., WPPSI-R, WISC-III, WAIS-III, referidas aqui nas suas edições mais recentes).

Todas as revisões dos principais instrumentos de avaliação psicológica colocam questões relativas à comparabilidade dos resultados com as versões anteriores. No caso da WISC-III, o termo de comparação mais imediato e habitual é a WISC-R. Entre nós, o termo de comparação da nova versão portuguesa da WISC-III é a anterior e única versão aferida disponível da WISC, adaptada por Ferreira Marques e utilizável desde 1970¹.

Uma das questões que importa analisar, em termos de investigação, remete para a relação entre os resultados obtidos na WISC-III e na WISC (a anterior e única versão disponível para a população portuguesa). A partir de agora vai ser possível, e frequente, os psicólogos fazerem reavaliações, com a WISC-III, de sujeitos previamente testados com a WISC. A comparabilidade entre os resultados alcançados nas versões sucessivas da WISC (WISC-III/WISC-R ou, no contexto português, WISC-III/WISC) constitui um requisito importante em situações de reavaliação. Os casos que precisam de continuar a receber apoio do ensino

especial, e necessitam de ser reexaminados com alguma periodicidade, ou as situações de monitorização de ganhos cognitivos resultantes da implementação de uma estratégia de intervenção constituem exemplos específicos de situações que requerem reavaliação com recurso a uma escala de inteligência.

No presente estudo procede-se à comparação e ao estudo das diferenças nos resultados entre a versão mais antiga da WISC e a nova versão (WISC-III). Este tipo de estudo tem um interesse particular quando se pretende fazer a “transposição” dos resultados na WISC para a WISC-III, isto é, conhecer o significado “actual” dos resultados obtidos com uma versão mais antiga. Uma observação “intuitiva” dirá que as normas de uma escala da inteligência como a versão portuguesa da WISC, aferida na década de 60, se encontra há muitos anos datada e a necessitar de uma revisão urgente. Os psicólogos portugueses reconhecem há muito tempo que a WISC sobreavalia a inteligência das crianças portuguesas. É este o cenário que explica actualmente o recurso à versão espanhola da WISC-R, em detrimento da versão portuguesa da WISC, por parte de alguns psicólogos portugueses. Do nosso ponto de vista, tal prática é discutível.

Por outro lado, os resultados desta investigação são importantes e têm igualmente utilidade prática uma vez que o tipo de casos abrangidos (crianças e adolescentes com dificuldades escolares e/ou problemas de comportamento) constitui uma parte muito substancial dos casos encaminhados para serviços de psicologia (avaliação e consulta psicológica) e são frequentemente examinados, várias vezes, durante o seu percurso escolar. Também neste contexto é compreensível a utilidade da comparação entre os resultados na

¹ Assim, entre a edição de ambas as escalas decorreram mais de 30 anos. Ainda que estes instrumentos tenham sido construídos segundo o mesmo modelo e numerosos itens sejam comuns, os resultados que eles permitem recolher são sensivelmente diferentes.

WISC (versão até hoje mais utilizada) e na WISC-III (que já pode ser empregue pelos psicólogos portugueses).

Este tipo de investigação alerta-nos ainda para as implicações decorrentes da disponibilização de uma aferição recente de uma nova versão da Escala de Inteligência de Wechsler para crianças ao nível do diagnóstico, encaminhamento, orientação e apoio especializados.

Considerando a actualização das normas na WISC-III e tomando exclusivamente como referência a evolução possível dos resultados obtidos, por um mesmo sujeito, anteriormente na WISC e, a partir de agora, na WISC-III, é plausível a ocorrência de alterações de natureza diagnóstica. É previsível que em muitos dos casos anteriormente testados com a WISC se verifiquem pontuações comparativamente mais baixas quando reavaliados com a WISC-III. Assim, e a título de exemplo, alguns casos previamente identificados, a partir da WISC, como tendo dificuldades cognitivas moderadas (com QIs compreendidos entre 70/75 – 80) passam a ser enquadrados, quando reavaliados com a WISC-III, no diagnóstico de Deficiência Mental (nível intelectual “Muito Inferior”). Na mesma ordem de ideias, é igualmente provável que algumas crianças e adolescentes considerados, a partir da WISC, como sobredotados do ponto de vista intelectual (QIs ≥ 120 ; classificação “Superior” ou “Muito Superior”) deixem de o ser na WISC-III (passando para um nível intelectual correspondente a uma classificação “Médio Superior” ou mesmo “Médio”).

Método

As versões portuguesas da WISC-III e da WISC foram administradas, de forma contra-balanceada no que respeita à ordem

de apresentação, a uma amostra de 70 crianças e adolescentes, observados em contexto clínico e com idades compreendidas entre os 6 e os 15 anos (mediana=10 anos; 67% do sexo masculino e 33% do sexo feminino). Clinicamente, este grupo é bastante heterogéneo: 23% dos casos apresentam dificuldades de aprendizagem, 70% correspondem a situações de co-morbilidade (dificuldades de aprendizagem e problemas de comportamento como hiperactividade com défice de atenção, oposição, ansiedade e/ou depressão), 3% referem-se a problemas de natureza familiar, 1% a casos de epilepsia e 3% a casos clínicos insuficientemente definidos. Todos os sujeitos foram examinados no PSIANIMA (Centro de Estimulação e Consulta Psicológica, em Espinho).

O intervalo entre as duas aplicações oscilou entre 14 e 89 dias (mediana do intervalo = 41 dias). A maioria das crianças e adolescentes provém de famílias de níveis socioeconómicos baixos, tratando-se, sobretudo, de filhos de trabalhadores não qualificados e, mais raramente, de trabalhadores da produção industrial.

Alguns sujeitos já tinham, anteriormente, respondido à WISC (mas não nos últimos 12 meses).

Apesar da precaução de administrar as provas numa ordem perfeitamente contra-balanceada, tal procedimento equilibra mas não elimina os efeitos da prática (Doll & Boren, 1993). Recorde-se que esses efeitos podem inflacionar os QIs numa grandeza desconhecida, bem como afectar a interpretação das diferenças QIV-QIR.

Resultados

No Quadro 1 são apresentadas as médias, desvios-padrão e correlações corrigidas

entre os resultados nos subtestes e QIs na WISC e WISC-III.

Neste estudo (cf. Quadro 1) as pontuações QIEC obtidas na WISC-III são, em média, aproximadamente, 28 pontos inferiores às

da linguagem). Estas comparações incidem sobre versões contíguas, sobretudo WISC-III/WISC-R e, mais raramente WISC-R/WISC. Ou seja, em nenhuma das investigações a seguir recenseadas se

QUADRO 1: Médias, desvios-padrão e correlações corrigidas entre os resultados nos subtestes e QIs na WISC-III e na WISC

Subteste/Escala	WISC		WISC-III		r_{12}
	Média	Desvio-Padrão	Média	Desvio-Padrão	
Informação	8.3	2.8	7.5	3.2	.61
Semelhanças	12.9	2.9	7.7	2.9	.47
Aritmética	9.0	2.4	8.3	2.2	.53
Vocabulário	10.7	2.2	7.4	2.8	.69
Compreensão	10.2	2.8	7.5	3.2	.59
Memória de Dígitos	9.0	2.3	8.2	3.0	.42
Completamento de Gravuras	10.9	2.5	8.0	3.4	.53
Código	13.1	3.4	7.8	3.2	.61
Disposição de Gravuras	13.9	3.0	8.3	3.5	.47
Cubos	12.9	3.1	8.4	2.6	.81
Composição de Objectos	12.9	2.4	8.2	3.2	.52
Labirintos	12.8	2.6	8.7	3.5	.58
Pesquisa de Símbolos	-	-	8.4	3.0	-
QI Verbal	101.4	12.6	83.4	14.9	.77
QI Realização	117.4	12.4	86.2	15.8	.74
QI Escala Completa	110.1	12.3	81.8	15.9	.78

Nota: N = 70. As correlações foram calculadas separadamente para cada ordem de administração, num desenho contra-balanceado, e foram corrigidas devido à variabilidade dos resultados da WISC-III (Guilford & Fruchter, 1978). A média ponderada foi obtida através da transformação Z de Fisher. As médias correspondem à média das médias das duas ordens de administração.

alcançadas na WISC. Os resultados no QIV e no QIR na WISC-III estão situados, respectivamente, 18 e 31 pontos abaixo dos resultados no QIV e no QIR da WISC. Estes resultados podem ser objecto de uma comparação aproximada com um conjunto numeroso de investigações que incluem as aferições americana e francesa/belga, amostras constituídas por sujeitos normais, amostras clínicas, grupos especiais (casos de sobredotação, dificuldades de aprendizagem, deficiência mental, problemas de comportamento, perturbações

procede à comparação por nós realizada entre a WISC-III e a WISC. Apesar de se tratarem de versões descontínuas, este é o único paralelismo que faz sentido entre nós, uma vez que a WISC-R nunca teve nenhum estudo de adaptação ou aferição para a população portuguesa.

Os manuais das aferições da WISC-III realizadas em diferentes países não constituem o melhor ponto de partida para interpretar os nossos resultados. Nestes manuais é possível encontrar os resultados de investigações sobretudo com sujeitos

considerados normais e centradas nas comparações WISC-III/WISC-R. Na investigação portuguesa a comparação é entre a WISC-III e a WISC e os sujeitos provêm de uma amostra clínica heterogénea. Sem nunca perder de vista estas diferenças, começemos por analisar os resultados das comparações presentes nalguns manuais.

Assim, e no âmbito da aferição norte-americana, a WISC-III (Wechsler, 1991) e a WISC-R (Wechsler, 1974) foram aplicadas a uma amostra de 206 sujeitos, 45% rapazes e 55% raparigas, com idades compreendidas entre os 6 e os 16 anos (média de idades = 11 anos). O intervalo médio das aplicações foi de 21 dias (amplitude de 12 a 70 dias). Em comparação com os resultados na WISC-R, e em média, na WISC-III o QIEC foi inferior em 5.3 pontos [108.2 (WISC-R) / 102.9 (WISC-III), com uma progressão média de 0.31 pontos por ano; 3.1 pontos por década]; o QIV é menor em 2.4 pontos [103.9 (WISC-R) / 101.5 (WISC-III), progressão média de 0.14 pontos] e o QIR é inferior em 7.4 pontos [111.6 (WISC-R) / 104.2 (WISC-III) progressão média de 0.44 pontos por ano] (cf. Wechsler, 1991, p. 89).

No contexto da aferição francesa/belga, a WISC-III (Wechsler, 1996) e a versão francesa da WISC-R (Wechsler, 1981) foram aplicadas a uma amostra de 99 sujeitos, 47% rapazes e 53% raparigas, com idades compreendidas entre os 6 e os 16 anos (média de idades = 11 anos e 7 meses). O intervalo médio das aplicações foi de 30 dias (amplitude de 19 a 64 dias). Em comparação com os resultados na WISC-R, e em média, na WISC-III o QIEC é inferior 3.7 pontos [107.9 (WISC-R) *versus* 104.2 (WISC-III), progressão média de 0.25 pontos por ano; 2.5 pontos por década]; o QIV é menor em 2.5 pontos

[progressão média de 0.17 pontos] e o QIR é inferior em 3.7 pontos [progressão média de 0.25 pontos por ano] (cf. Wechsler, 1996, p. 210).

Este padrão de resultados superiores na WISC-R, em relação aos obtidos na WISC-III pelos mesmos sujeitos, é encontrado de forma sistemática na literatura científica, independentemente da natureza das amostras investigadas (sujeitos sem problemas, sujeitos pertencentes a grupos clínicos). Resultados consistentemente mais elevados são também observados quando se comparam os desempenhos dos mesmos sujeitos na WISC e na WISC-R. Ou seja, a superioridade dos desempenhos nas versões mais antigas constitui a regra. Por exemplo, no confronto entre os desempenhos na WISC-R e na WISC foram observadas diferenças de 5 a 9 pontos em 58 crianças sem problemas (Schwartzing, 1976), ou diminuições estatisticamente significativas de 6.76 pontos (83.94-77.18) no QIEC, 2.74 pontos (78.42-75.68) no QIV, 9.42 pontos (92.36-82.94) no QIR da administração da WISC para a WISC-R (Vercoustre, 1985).

Ainda que incidindo nas comparações WISC-III/WISC-R, as investigações realizadas com amostras clínicas (Dificuldades de Aprendizagem e/ou Problemas de Comportamento) são mais pertinentes para interpretar os nossos resultados. No contexto português, muitas das crianças já examinadas com a WISC serão no futuro reavaliadas com a WISC-III. Daí, o interesse adicional por este tipo de estudos comparativos. Vejamos em primeiro lugar os resultados provenientes de amostras clínicas heterogéneas.

Numa revisão de vários estudos que abrangem 1000 sujeitos encaminhados para consultas de psicologia, e que foram objecto de reavaliação com a WISC-III,

9 meses ou mais depois, Zimmerman e Woo-Sam (1996) constataram a existência de diminuições estatisticamente significativas de 6.4 pontos (99.8-93.4) no QIEC, 5.60 pontos (99.3-93.7) no QIV, 5.5 pontos (100.0-94.5) no QIR da WISC-III.

Noutra pesquisa, as versões norte-americanas da WISC-III e da WISC-R foram aplicadas a uma amostra clínica, constituída por casos com dificuldades de aprendizagem (57%), déficit de atenção com hiperactividade (35%) e ansiedade ou depressão (8%). Os 104 sujeitos desta amostra tinham idades compreendidas entre os 7 e os 14 anos (e média de idades de 10 anos). O intervalo de tempo entre as duas aplicações foi de 12 a 70 dias (intervalo médio de 21 dias). Em comparação com os resultados na WISC-R, e em média, na WISC-III o QIEC é inferior em 5.9 pontos (progressão média de 0.37 pontos por ano); o QIV é menor em 5.4 pontos (progressão média de 0.34 pontos) e o QIR é inferior em 5.1 pontos (progressão média de 0.32 ponto por ano) (Wechsler, 1992, pp. 91-92).

Por sua vez, Graf e Hinton (1994) observaram uma amostra heterogénea constituída por 84 alunos com dificuldades de aprendizagem e/ou problemas de comportamento, cujos QIEC na WISC-III se situam entre 65 e 135 (média 95) e na WISC-R igualmente entre 60-135 (média 97). Após um intervalo de 3 anos os QIEC e (os QIR) são significativamente mais reduzidos na WISC-III comparativamente à WISC-R.

Uma outra investigação com a WISC e a WISC-R comparou os desempenhos de 46 crianças com problemas de comportamento ou suspeitas de dificuldades de aprendizagem, e constatou diferenças de 8 pontos (89.7-81.7) no QIEC, 7 pontos (89.3-82.3) no QIV, 8.3 (92.2-83.9) no QIR (Weiner & Kaufman, 1979).

Estas diferenças de 5 a 8 pontos são comumente encontradas noutras investigações (Doll & Boren, 1993; Newby, Recht, Caldwell, & Schafer, 1993; Slate, 1995), embora haja alguma indicação de que as discrepâncias possam ser maiores nalguns grupos de crianças excepcionais (Dumont & Faro, 1993; Post, 1992). Esta situação pode ter um impacto substancial em decisões de encaminhamento para instituições ou serviços de educação especial (Post & Mitchell, 1993).

Algumas pesquisas centram-se mais especificamente em grupos com dificuldades de aprendizagem (DA). Neste contexto, Lyon (1995) estudou um grupo de 40 sujeitos com DA que após uma avaliação inicial com a WISC-R foram objecto de reavaliação, 1.5 a 3.7 anos depois, com a WISC-III, tendo observado nesta última escala diminuições estatisticamente significativas de 7.05 pontos (98.58-91.53) no QIEC, 5.60 pontos (95.08-89.48) no QIV, 7.62 pontos (103.10-95.48) no QIR. Neste estudo a WISC-III foi sempre aplicada em segundo lugar (contrariando, a desejável ordem de aplicação contra-balanceada). No mesmo sentido, os resultados sistematicamente mais reduzidos em todos os QIs na WISC-III foram igualmente registados numa outra amostra de DA (Gunter, Sapp & Greene, 1995). A única pesquisa que conhecemos onde se observa um resultado superior na WISC-III é a de Kogan (1996). Esta investigadora analisou os desempenhos de 40 alunos com DA e uma média de idades de 10 anos e 9 meses, e constatou que o QIV obtido a partir da WISC-III foi superior ao alcançado por este grupo na WISC-R. De acordo com os resultados de todas as outras investigações, as pontuações no QIEC e no QIR foram mais baixas na WISC-III. No entanto, um exemplo das consequências da investigação em alunos

com DA é evidenciado na pesquisa de Scott (1999). Esta autora observou uma diminuição geral dos resultados, mais acentuada nos alunos afro-americanos e ainda que várias crianças anteriormente diagnosticadas com DA (a partir dos resultados na WISC-R) foram reclassificadas posteriormente com Deficiência Mental (com base nos retestes feitos com a WISC-III).

Numa amostra heterogênea de 61 crianças examinadas no contexto de reavaliações trienais e que frequentavam instituições ou serviços de educação especial, o QIEC da WISC-III foi 5 a 9 pontos inferior ao QIEC da WISC-R (Aichinger, Bolen, Hall & Webster, 1993). Noutras pesquisas, a diferença de resultados no QIEC entre amostras de alunos com diferentes condições deficitárias foi de 0.9 (Klein & Fisher, 1993), 5.69 (Smith, Stoval & Geraghty, 1994), 7.95 (Bolen, Aichinger, Hall & Webster, 1995) e 8 pontos (Jones, 1996). Mais recentemente, Carlton e Sapp (1997) examinaram uma amostra de 65 sujeitos (40 rapazes e 17 raparigas, com uma média de idades de 12 anos e 7 meses), que recebiam serviços de ensino especial, inicialmente testados com a WISC-R e avaliados 3 anos depois com a WISC-III. Esta amostra incluía casos de DA (83%), perturbações do comportamento (9%) e deficiência mental (8%). Foram assinaladas diminuições estatisticamente significativas de 4.09 pontos (91.54-87.45) no QIEC, 3.12 pontos (92.15-89.03) no QIV e 4.44 pontos (92.54-88.10) no QIR. Numa amostra de 93 crianças com deficiência mental moderada, submetidas a avaliações trienais, os QIEC da WISC-III são 5 a 8 pontos inferiores ao QIEC da WISC-R (Nagle & Daley, 1994). QIs 6 a 9 pontos mais reduzidos na WISC-R (relativamente à WISC) foram igualmente observados numa amostra de

48 crianças com atraso mental educável (Hamm *et al.*, 1976). Este tipo de dados contraria a “clarificação das conclusões de Flynn” apresentada por Spitz (1989). Com base em dados comparativos entre resultados na WAIS e WAIS-R este investigador assinala que os ganhos de 3 pontos por década são verdadeiros para as pessoas com QIs médios, mas perdem importância (isto é, diminuem gradualmente) à medida que os desempenhos se afastam da amplitude média e podendo mesmo a direcção dos resultados ser inversa nos sujeitos com atraso mental.

Nos alunos sobredotados e talentosos constata-se igualmente, e de forma sistemática, resultados inferiores na WISC-III. As diferenças nos QIEC oscilam entre 1.35 (Gridley, Arceneaux, Gayer, Buelow & Palmer, 1994), 12.83 (Sevier & Bain, 1994)² e 18 pontos (Bryant, 1992). O estudo de Larrabee e Holroyd (1976) que comparou os desempenhos de 38 crianças “intelectualmente brilhantes” na WISC-R e na WISC refere também resultados 8 a 10 pontos mais baixos na WISC-R.

No mesmo sentido vai a investigação de Doll e Boren (1993) que examinaram 25 crianças com perturbações da linguagem, com idades compreendidas entre os 8 aos 13 anos (mediana 9 anos e meio) e observaram diminuições estatisticamente significativas de 6 pontos (83.17-77.20) no QIEC, 5.4 (79.13-73.72) pontos no QIV e 5.6 pontos (90.13-84.52) no QIR.

² Nesta investigação é possível constatar que as diferenças relativas ao QIV (14.57 pontos) são superiores às discrepâncias no QIR (7.60). Além disso, enquanto na WISC-R os resultados padronizados oscilavam entre 15 e 19 pontos na maioria dos subtestes, na WISC-III esse valor alternava sobretudo entre os 12 e 15 pontos padronizados.

Por outro lado, a investigação de Bolen, Aichinger, Hall e Webster (1995) introduz talvez um elemento adicional para interpretar a magnitude dos nossos resultados. Bolen e colaboradores constataram diferenças de aproximadamente 8 pontos nos QIEC (e de 9 pontos nos QIR) nas comparações quer entre a WISC-R e a WISC-III, quer entre a WISC e a WISC-R, a partir de uma amostra de 61 alunos com problemas de aprendizagem e dificuldades de natureza cognitiva que por isso frequentavam serviços de educação especial.

As correlações elevadas entre os QIs Verbal (.77), de Realização (.74) e da Escala Completa (.78) de ambas as Escalas (cf. Quadro 1) constituem uma confirmação de que a WISC-III mede essencialmente os mesmos constructos que a WISC.

Nos estudos publicados no âmbito de outras aferições da WISC-III, as correlações entre os QIs (QIV, QIR, QIEC) da WISC-III e da WISC-R são um pouco mais substanciais com sujeitos normais: .90 (QIV), .81 (QIR), (.89) QIEC (Wechsler, 1991); .89 (QIV), .80 (QIR), (.88) QIEC (Wechsler, 1996). Estas correlações são ligeiramente mais reduzidas, mas ainda assim elevadas, em amostras clínicas: .88 (QIV), .79 (QIR), .88 (QIEC) (Wechsler, 1991); .86 (QIV), .82 (QIR) e .88 (QIEC) (Carlton & Sapp, 1997); .82 (QIV), .81 (QIR) e .86 (QIEC) (Graf & Hinton, 1994). Várias investigações utilizando as escalas WISC e WISC-R evidenciam que são igualmente relevantes (.85) as correlações entre QIEC (Hamm, Wheeler, McCallum, Herrin, Hunter & Catoe, 1976; Larrabee & Holroyd, 1976).

Quando a análise dos resultados se desloca dos QIs para os subtestes é compreensível que o valor das correlações diminua.

Assim, na nossa investigação as correlações entre os subtestes das duas escalas (WISC/WISC-III), oscilam entre .42 (Memória de Dígitos) e .81 (Cubos) (mediana = .57) (cf. Quadro 1).

Na aferição americana as correlações ao nível dos subtestes variaram entre .42 (Complemento de Gravuras) e .80 (Informação) (mediana = .66) (Wechsler, 1991). Na aferição francesa/belga, os valores de correlação situaram-se entre .43 (Composição de Objectos) e .79 (Informação) (mediana = .62) (Wechsler, 1996). Numa amostra de 46 crianças afro-americanas e com suspeita de DA e problemas de comportamento, Weiner e Kaufman (1979) obtiveram correlações entre os subtestes da WISC e da WISC-R que variaram entre .51 (Complemento de Gravuras) e .91 (Informação) (mediana = .69).

No nosso estudo, a Informação, a Aritmética e a Memória de Dígitos constituem os únicos subtestes da WISC em que as pontuações se encontram abaixo da média normativa de 10 pontos. Curiosamente estes são três dos subtestes que constituem o perfil ACID (que inclui os subteste de Aritmética, Código, Informação e Dígitos) e que para alguns investigadores seria característico dos disléxicos (Kaufman, 1981; Spafford, 1989). No entanto, as pontuações na WISC-III não corroboram a pertinência deste perfil, nem no nosso estudo (onde uma percentagem muito elevada de casos apresenta DA isoladamente ou associadas a outros problemas clínicos), nem na pesquisa de Prifitera e Dersh (1993) em que apenas 5.1% dos casos de DA e 12.3% dos casos com perturbações da atenção manifestam resultados deficitários naquelas provas. Ainda na WISC, os resultados são mais elevados na maior parte dos subtestes de realização (com excepção do subteste de Complemento de Gravuras). A

pontuação mais elevada na parte verbal da WISC ocorre no subteste de Semelhanças (prova em que é, aliás, consensual, por parte dos psicólogos portugueses, que as normas estão há muito ultrapassadas). Pelo contrário, na WISC-III as pontuações em todos os subtestes (verbais e de realização) são homoganeamente inferiores à média (os valores, entre 7.4 e 8.7, apresentam uma variabilidade bastante reduzida e inviabilizam qualquer tentativa de definição de um perfil característico). Pelo contrário, na WISC-III as pontuações em todos os subtestes (verbais e de realização) são homoganeamente inferiores à média: os valores observados, entre 7.4 e 8.7, apresentam uma variabilidade bastante reduzida.

Ainda no Quadro 1 é possível observar que na WISC existe uma maior amplitude e variabilidade dos resultados, bem evidente na diferença de 16 pontos entre o Quociente Intelectual Verbal (QIV) e o Quociente Intelectual de Realização (QIR). Na WISC-III essa discrepância é de apenas 2.8 pontos. O perfil QIR>QIV, que é encontrado com frequência em sujeitos com DA (Prifitera

máximos) e diferenças entre os QIs na WISC e na WISC-III. São estatisticamente significativas ($p > .01$) as diferenças entre as médias dos QIV, QIR e QIEC da WISC e da WISC-III. Os valores mais elevados são encontrados nos QIs da WISC. Os resultados mais reduzidos são observados nos QIs da WISC-III.

O Quadro 3 mostra as médias do QIV, QIR, QIEC na WISC-III e na WISC, tomando como referência três níveis de QIEC na WISC-III. Pode observar-se, uma vez mais, que os resultados são sistematicamente superiores na WISC, independentemente do nível intelectual considerado. Constata-se igualmente que as diferenças QIV-QIR são sempre mais acentuadas na WISC (15.4 a 17.2) do que na WISC-III (0.9 a 3.8), qualquer que seja o nível intelectual analisado.

No Quadro 4 estão assinaladas as diferenças QIV-QIR na WISC e na WISC-III. Pode constatar-se que as discrepâncias são em maior número na WISC do que na WISC-III, qualquer que seja o critério usado (≥ 10 , >12 , >15 pontos). Há também um número esmagador de casos na WISC

QUADRO 2: Médias, Desvios-Padrão, Amplitudes e Diferenças entre os QIs da WISC e da WISC-III

Escala	WISC			WISC-III			/
	M	DP	Min-Máx	M	DP	Min-Máx	
QIV	101.4	12.6	66-128	83.4	14.9	53-118	16.2**
QIR	117.4	12.4	87-147	86.2	15.8	55-135	21.5**
QIEC	110.1	12.3	75-134	81.8	15.9	52-120	22.9**

** $p > .01$

& Dersh, 1993; Wechsler, 1996), é bem menos evidente na nossa investigação (cf. igualmente Quadros 3 e 4).

O Quadro 2 apresenta as médias, desvios-padrão, amplitudes (valores mínimos e

em que o QIR é superior ao QIV (66 em 70 casos). Embora não seja a situação mais frequente, em termos comparativos a discrepância QIV>QIR é bem mais acentuada na WISC-III (24 casos) do que na WISC (3 casos).

QUADRO 3: Médias do QIV, QIR, QIEC na WISC-III e na WISC por nível de QIEC na WISC-III

	QI _{EC} : 40-69 (N=18)	QI _{EC} : 70-89 (N=30)	QI _{EC} : 90 (N=22)
WISC-III			
QIV	67.4	80.3	100.7
QIR	68.3	84.1	103.7
QIEC	63.4	78.5	101.3
WISC			
QIV	89.8	101.4	110.9
QIR	107.0	117.1	126.3
QIEC	98.2	110.0	120.1

Discussão

Convém insistir no facto da presente investigação apenas permitir a comparação entre os resultados com as versões portuguesas da WISC e da WISC-III. O confronto dos resultados das investigações realizadas noutros países processa-se habitualmente entre a WISC e a WISC-R ou, mais frequentemente, entre a WISC-R e a WISC-III. Por outras palavras, neste segundo conjunto de investigações os resultados são cotejados entre versões sucessivas (WISC/WISC-R ou WISC-R/WISC-III) e não, como é o caso do presente estudo português, entre versões descontínuas (WISC/WISC-III).

Os testes são diferentes e a continuidade estrutural é compreensivelmente menor quando se compara a WISC-III com a WISC. Alguns posicionamentos conhecidos, relativamente ao confronto entre a WISC-III e a WISC-R, não são necessariamente aplicáveis à comparação, para nós mais pertinente, entre a WISC-III e a WISC. A este respeito Dixon e Anderson (1995) mostram que as estruturas de covariância das amostras normativas não diferem para as duas escalas. E, nesta mesma linha, Teeter e Korducki (1998) defendem que as duas últimas versões da

WISC (WISC-III e WISC-R) são suficientemente similares pelo que a investigação com a WISC-R pode ser estendida à WISC-III³.

Por outro lado, e ainda que todos estes instrumentos tenham sido construídos segundo o mesmo modelo, e numerosos itens sejam idênticos ou iguais, o denominador comum é maior entre versões sucessivas, isto é, entre a WISC-R e a WISC-III (ou entre a WISC-R e a WISC) e menor entre a WISC e a WISC-III. Além disso, importa destacar desde já que não estão em causa as qualidades psicométricas dos dois instrumentos (WISC e WISC-III). A sua fiabilidade é quase idêntica e as suas normas são igualmente representativas da população portuguesa, entre os 6 e os 15 anos, ou seja, as idades comuns de aplicação das duas provas (Ferreira Marques, 1969; Wechsler, 2003).

É necessário sublinhar, que uma parte da explicação para a diferença entre os

³ Existem obviamente algumas diferenças entre as duas escalas como é, por exemplo, o caso, na WISC-III, da possibilidade de cálculo de Índices factoriais, a presença de um novo subteste (Pesquisa de Símbolos) e, aparentemente, um maior peso da rapidez de desempenho nos subtestes de realização.

QUADRO 4: Diferenças QIV-QIR na WISC e na WISC-III

WISC	
DIF. 10 PONTOS:	52 casos (74.3 %)
DIF. > 12 PONTOS:	43 casos (61.4 %)
DIF. > 15 PONTOS:	34 casos (48.6 %)
WISC-III	
DIF. 10 PONTOS:	23 casos (32.9 %)
DIF. > 12 PONTOS:	16 casos (23.0 %)
DIF. > 15 PONTOS:	12 casos (17.1 %)
WISC	
QIV > QIR:	3 casos (4.3%)
DIF. 10 PONTOS:	0 casos (0 %)
DIF. > 12 PONTOS:	0 casos (0 %)
DIF. > 15 PONTOS:	0 casos (0 %)
QIR > QIV:	66 casos (94.3%)
DIF. 10 PONTOS:	52 casos (74.3 %)
DIF. > 12 PONTOS:	43 casos (61.4 %)
DIF. > 15 PONTOS:	36 casos (51.4 %)
QIV = QIR:	1 caso (1.4%)
WISC-III	
QIV > QIR:	24 casos (34.3%)
DIF. 10 PONTOS:	5 casos (7.1 %)
DIF. > 12 PONTOS:	4 casos (5.7 %)
DIF. > 15 PONTOS:	2 casos (2.9 %)
QIR > QIV:	41 casos (58.6%)
DIF. 10 PONTOS:	18 casos (25.7 %)
DIF. > 12 PONTOS:	12 casos (17.1 %)
DIF. > 15 PONTOS:	10 casos (14.3 %)
QIV = QIR:	5 casos (7.1%)

resultados obtidos na WISC e na WISC-III resulta de importantes transformações históricas, políticas, sociais e educativas ocorridas no nosso país, entre a década de 60 e os nossos dias (distância temporal entre as duas aferições). Mais especificamente, entre a edição da versão

portuguesa da WISC-III (Wechsler, 2003) e a da primeira versão portuguesa da WISC (Ferreira Marques, 1969) decorreram 34 anos. No entanto, o intervalo entre a recolha dos dados nas duas versões é maior, oscilando entre 33 e 41 anos (cf. Wechsler, 2003; Ferreira Marques, 1969, páginas 139

e 141). Em 30/40 anos as características (educativas, socioeconómicas, etc.) da população portuguesa evoluíram. Neste período de tempo é de supor que as características intelectuais da população portuguesa tenham sofrido alterações, e que o QI médio das crianças portuguesas tenha progredido. Ou seja, à semelhança do que aconteceu noutros países, aumentaram os desempenhos médios das crianças e adolescentes portugueses numa mesma prova de inteligência (a WISC, no caso em apreço).

Finalmente, recorde-se que o grupo estudado é bastante heterogéneo do ponto de vista clínico e 70% dos sujeitos observados correspondem mesmo a casos de co-morbilidade. No seu conjunto, e isoladamente, estes aspectos não podem ser esquecidos na interpretação dos nossos dados. Eles podem ajudar a explicar, pelo menos em parte, o facto da divergência dos resultados ser sensivelmente maior no nosso estudo.

Esta evolução nos resultados pode surpreender. No entanto, trata-se de um fenómeno robusto que tem sido observado na maior parte dos países ocidentais, depois dos anos trinta (Flynn, 1987). Este fenómeno é frequentemente designado por *efeito de Flynn*, adoptando o nome do investigador neo-zelandês, James R. Flynn, que foi o primeiro a observá-lo. Este investigador constatou que entre 1932 e 1972, os resultados dos norte-americanos em testes normalizados tinham aumentado 0.3 pontos de QI por ano (Flynn, 1984), cerca de 3 pontos por década (Flynn, 1987) e que esta progressão era similar noutros países ocidentais (Flynn, 1987). Este tipo de ganhos nos QIs foi identificado em mais de 20 países: Estados-Unidos, Holanda, Bélgica, França, Noruega, Suécia, Dinamarca, Alemanha, Áustria, Suíça,

Escócia, Irlanda do Norte, Canadá, Austrália, Nova Zelândia, Brasil, Japão e áreas urbanas da China (Flynn, 1998). Flynn (*in press*) especifica que no espaço de uma geração (30 anos), nas subescalas verbais das Escalas de Inteligência de Wechsler os ganhos foram de aproximadamente 9 pontos. Este progresso é ainda mais elevado para os testes de inteligência não verbais. Nas subescalas de realização das Escalas de Inteligência de Wechsler os resultados oscilam entre os 9 e os 20 pontos. Em vários países ocidentais (Áustria, França, Bélgica, Holanda, Japão), a evolução dos desempenhos nas Matrizes Progressivas de Raven é de um aumento de 6 a 8 pontos de QI, por década (cf. Flynn, 1987). Na Alemanha, os ganhos médios são de 10 pontos por década (Kaufman, 1990). Actualmente, as crianças, adolescentes (e adultos) respondem correctamente a um maior número de questões verbais e resolvem mais problemas não verbais do que há 5, 10, 20 ou 30 anos (Flynn, 1984, 1987; Kaufman, 1990).

Quais as causas ou explicações para estes ganhos no desempenho em testes de inteligência? Têm sido avançadas várias explicações para este fenómeno cuja existência é actualmente indiscutível (Flieller, 2001). É possível que uma parte importante desta evolução seja determinada por variáveis como a melhoria qualitativa e quantitativa da dieta alimentar e dos cuidados de saúde, maior complexidade ambiental com o conseqüente aumento das exigências cognitivas do meio, o prolongamento da duração dos estudos (incluindo a escolarização mais precoce das crianças e, no caso português, a melhoria das condições de vida e a democratização no acesso ao ensino), a natureza dos jogos educativos presentemente propostos e a

proliferação de material figurativo, o aumento do nível de instrução dos pais, a diminuição do número de crianças por família, crianças mais informadas, etc. (Kaufman, 1990; Mackintosh, 1998). Todavia, estes argumentos parecem conter um certo número de insuficiências, quando considerados isoladamente. Outras variáveis por vezes referidas, como é o caso da prática na resposta a testes, as mudanças de atitude ou de estratégia e a elevação da qualificação profissional dos pais apresentam um valor explicativo negligenciável para este *efeito de Flynn*. Por exemplo, segundo Flieller (2001) os ganhos de qualificação profissional dos pais, não permitem explicar mais do que 5% da progressão dos resultados nos testes.

Esta evolução, bem como a respectiva magnitude, do nível médio de QI no seio da população contém importantes repercussões práticas e nem sempre é detectada pelos psicólogos. Frequentemente, os psicólogos não têm consciência de que as transformações contínuas das características cognitivas da população obrigam a re-aferir, periodicamente, os testes de inteligência e persistem em aplicar testes cujas normas estão obsoletas. Esta prática apenas pode ser compreensível quando não existam alternativas válidas e desde que o valor quantitativo dos resultados seja relativizado.

Mais discutível é a insistência na utilização de uma escala com normas muito datadas apesar de existirem provas alternativas, mais recentes e com normas actualizadas. Neste contexto, sabe-se que alguns psicólogos manifestam por vezes alguma “relutância” ou “resistência” em usar desde logo, aquando da sua publicação, uma nova versão das escalas de inteligência de Wechsler. Este cenário tem sido observado

em vários países. Alguns psicólogos que trabalham com crianças e adolescentes não têm em conta esta evolução do nível médio do QI na população e acusam a nova versão do teste (WISC-R ou WISC-III) de ser muito exigente e optam, de forma deliberada, por continuar a usar a versão anterior da Escala de Inteligência de Wechsler para Crianças (WISC ou WISC-R, respectivamente). Mais especificamente, Kaufman (1994) lembra que, quando a WISC-R apareceu nos Estados-Unidos, alguns psicólogos manifestaram inquietação e o sentimento de estarem a fazer “algo de errado”, por esta nova escala produzir QIs mais baixos em 5 ou 10 pontos relativamente à versão anterior (WISC). Uma outra investigação realizada por Harrison, Kaufman, Hickman & Kaufman (1988) mostra, igualmente, que alguns anos após o aparecimento da WAIS-R havia ainda um número importante de psicólogos que, pelos mesmos motivos, recorria à versão anterior (WAIS). Finalmente, na pesquisa desenvolvida por Herring e Reitan (1986) é também referido que as escalas Wechsler-Bellevue ainda eram utilizadas em neuropsicologia meio século depois do seu aparecimento, apesar da existência de novas versões, entretanto disponibilizadas (WAIS, WAIS-R). Não surpreende, por isso, que alguns destes profissionais considerem, por vezes, as novas versões dos testes demasiado exigentes e, por esta razão, continuem a utilizar a versão anterior disponível da escala de inteligência de Wechsler. Este cenário (preferência pela versão anterior devido ao decréscimo dos resultados com a nova versão) está igualmente documentado, em relação à WISC-R, em países como os Estados-Unidos ou a França (Wechsler, 1991; Wechsler, 1996). Convém ter consciência que uma tal prática é irresponsável, pois conduz a sobreavali-

os sujeitos testados em relação aos sujeitos da população à qual pertencem. Este procedimento pode conduzir a decisões inadequadas, prejudiciais aos sujeitos avaliados. No caso português, pretender eventualmente continuar a usar a versão mais antiga da WISC em detrimento da nova versão (WISC-III), significa comparar os desempenhos das crianças e adolescentes de hoje aos dos sujeitos da mesma faixa etária de há 30/40 anos. Como refere Grégoire (1992) o psicólogo deve recordar-se que testar é sempre comparar. Mas este tipo de comparação não é válido e constitui um erro metodológico e uma falha deontológica (American Psychological Association, 1999). Por consequência, é imperativo utilizar sempre testes cuja aferição seja a mais recente possível. No nosso caso, os psicólogos devem abandonar o uso da WISC, que inflaciona consideravelmente as pontuações, e passar a usar a versão portuguesa da WISC-III.

Parece-nos muito importante voltar a sublinhar que, na maior parte dos estudos que acabamos de recensear, as investigações seguem um desenho distinto do utilizado por nós. Nestas investigações, a WISC-III foi aplicada sistematicamente depois da WISC-R, e não numa ordem contra-balanceada. Estes estudos são realizados em contextos de avaliação psicológica clínica ou educacional, em que os sujeitos foram examinados num primeiro momento com a versão da WISC mais recente (WISC-R) e, meses ou anos depois, voltaram a ser re-examinados com a WISC-III, entretanto editada. Daqui decorrem inevitáveis efeitos da prática e de aprendizagem (que influenciam, em maior ou menor grau, mas sempre no mesmo sentido, as respostas dadas na WISC-III e que amenizam, forçosamente, as

diferenças entre a WISC-R e a WISC-III). Estas diferenças seriam provavelmente maiores se o desenho da maior parte das referidas investigações tivesse contemplado, como na nossa pesquisa, uma ordem de administração contra-balanceada. Este aspecto é muitas vezes ignorado na literatura consultada.

Independentemente dos factores causais, os resultados consideravelmente mais baixos numa nova versão (p. ex., na WISC-III relativamente à WISC-R, ou na WISC-R comparativamente à WISC) colocam sempre problemas nas re-avaliações, nomeadamente ao nível das implicações relativas à decisão de natureza diagnóstica e de orientação ou encaminhamento. O mesmo se passa, num grau compreensivelmente mais acentuado, com a nova versão portuguesa da WISC-III (Wechsler, 2003), quando os resultados com ela obtidos são comparados com os resultados da primeira e única versão disponível em Portugal (WISC, Ferreira Marques, 1969).

A questão que se coloca releva, pois, da possibilidade da modificação nas pontuações alterar diagnósticos em contextos de educação especial invertendo decisões relativas à orientação, encaminhamento e orientação. Neste contexto, é provável que com a introdução da WISC-III alguns casos com diagnóstico de dificuldades de aprendizagem⁴ deixem de justificar essa classificação e, de forma mais problemática, sejam excluídos ou percam a possibilidade de acesso a programas de apoio especial, ou de elegibilidade para outros recursos ou serviços apropriados. Neste contexto, é

⁴ De modo controverso, o diagnóstico de dificuldades de aprendizagem está fortemente assente na discrepância entre os desempenhos escolares deficitários e a normalidade da inteligência.

também plausível que a diminuição dos valores de QI legitime posturas de demissão por parte de alguns educadores, professores e outros profissionais que, assim, deixam de mobilizar esforços de natureza remediativa.

Sabe-se que ocorre uma diminuição sistemática dos resultados, quando a WISC-III é administrada depois de uma primeira aplicação da WISC-R, alterando em muitos casos o diagnóstico inicialmente formulado. Na pesquisa de Simione (1999), realizada com 190 crianças e adolescentes dos 6 aos 16 anos, 24% dos diagnósticos (de DA e outros) foram modificados em consequência da evolução dos resultados da WISC-R para a WISC-III. Por outro lado, algumas das incorrecções ou dos erros na classificação diagnóstica podem resultar da escolha dos testes utilizados, que constituem por esse motivo, eles próprios, um problema inerente ao sistema diagnóstico actual (Slate *et al.*, 1997). Por isso, deve ser colocado um cuidado particular tanto na interpretação qualitativa das diferenças (diminuição) de resultados da WISC para a WISC-III, como nas eventuais alterações na classificação diagnóstica do nível intelectual.

Os psicólogos (e outros profissionais) devem estar conscientes deste cenário, interrogar-se acerca do valor científico dos critérios de diagnóstico (re)conhecendo, nomeadamente, as objecções levantadas ao uso (ou peso excessivo) do QI como critério de diagnóstico, não negligenciando, em particular, os valores do erro padrão da medida na interpretação dos QIs. Ao mesmo tempo, os psicólogos devem aprofundar e diversificar a avaliação, recorrendo a outros instrumentos ou tarefas, com o objectivo de encontrar respostas que viabilizem intervenções mais eficazes.

Os resultados desta investigação sugerem a necessidade de um trabalho de ajustamento “mental”, por parte dos psicólogos, quando estes comparam o QI observado numa criança na WISC-III, com o QI anteriormente obtido pela mesma criança na WISC. Neste contexto, os psicólogos devem ser cautelosos antes de inferir uma perda no funcionamento cognitivo uma vez que QIs mais baixos são esperados quando se recorre à versão mais recente. Ou seja, a modificação das pontuações que resulta do processo de revisão e aferição dos instrumentos deve ser considerada como algo de normal. Vários investigadores assinalam explicitamente esta modificação. Assim, Reynolds (1990) sublinha que a diminuição dos resultados é um fenómeno esperado quando os instrumentos são reaferridos. Sattler (1992) escreve que “os sujeitos apresentam quase sempre desempenhos mais baixos nos novos testes do que nos testes mais antigos” (p. 1041). E Kaufman (1994) acrescenta que a introdução de normas actualizadas de uma nova prova torna visível uma diminuição das pontuações obtidas em comparação com as pontuações definidas pelas normas antigas. Neste contexto, Weiner e Kaufman (1979) referem que mesmo uma discrepância de 1 desvio-padrão (15 pontos) entre os resultados na WISC e na WISC-R pode não ser significativa.

Este tipo de pesquisa que compara os resultados em duas versões distintas de uma escala de inteligência (no nosso caso, WISC e WISC-III) levanta ainda um outro tipo de questão. O exame da inteligência é realizado no pressuposto de que a inteligência é um constructo fortemente estável. Este pressuposto tem recebido apoio considerável dos resultados de investigações realizadas na população em

geral (Stavrou, 1990; Wechsler, 1991), mas está menos bem documentado nas crianças e adolescentes excepcionais. Admite-se que estas crianças possam evidenciar uma variabilidade significativa nos seus desempenhos e apresentar “perfis erráticos” nos testes (Bolen, Aichinger, Hall & Webster, 1995; Webster, 1988). A partir de um estudo longitudinal com a duração de seis anos Stavrou (1990) refere que o QIEC constitui, a longo prazo, a medida ou indicador mais estável nas crianças excepcionais, mas chama a atenção para as frequentes alterações, clinicamente significativas, nos resultados dos sujeitos individualmente considerados e nas correlações teste-reteste (que são mais reduzidas nestas amostras do que nos grupos de crianças consideradas normais)⁵. Do nosso ponto de vista, os resultados desta investigação, e das pesquisas aqui recenseadas, não contrariam de forma evidente este pressuposto da estabilidade da inteligência, pois que assentam em comparações com versões distintas de escalas de inteligência, que dispõem de normas igualmente diferenciadas. O decréscimo dos resultados nas investigações que comparam sistematicamente os resultados de uma primeira aplicação com a WISC-R com os obtidos numa posterior administração aos mesmos sujeitos da WISC-III corresponde, como

⁵ Na linha dos estudos teste-reteste, e com uma postura mais crítica, Cassidy (1998) interroga-se mesmo acerca da necessidade (e utilidade) das reavaliações trienais e questiona as aplicações de rotina da WISC-III a todas as crianças envolvidas em contextos de educação especial. Com base nos ficheiros de 592 crianças esta investigadora constatou que muitas crianças com diagnósticos de deficiência mental, ou de perturbações do comportamento, apresentavam flutuações significativas nos seus resultados aquando das reavaliações trienais com a WISC-III. A mesma instabilidade foi comprovada em crianças com QIs superiores a 109.

assinalam Slate *et al.* (1997), a um “artefacto das normas e não uma diminuição do nível intelectual” (p. 120).

Neste momento já é possível usar em Portugal a WISC-III (Wechsler, 2003) e comparar os desempenhos de uma criança anteriormente examinada, com o auxílio da WISC (Ferreira Marques, 1970). A versão portuguesa da WISC-III permite chamar a atenção para os valores dos QIs na WISC estarem sobrestimados ou inflacionados por normas excessivamente datadas e que, obviamente, não são representativas dos desempenhos de crianças e adolescentes dos nossos dias. Neste contexto, Flynn (*in press*) defende que os testes de QI necessitam de ser normalizados de sete em sete anos, caso se pretenda que eles continuem a desempenhar um papel minimamente respeitável. Caso contrário, adverte ainda Flynn, os QI ficarão inflacionados, uma vez que os desempenhos serão inevitavelmente comparados com os de uma amostra de aferição, fora do prazo de validade.

A dimensão normativa da avaliação ocupa neste plano um papel decisivo. Kaufman (1993) defende mesmo que “as normas são o nome principal do jogo” (da avaliação com testes). Convém novamente recordar que as normas da primeira versão portuguesa da WISC (Ferreira Marques, 1969) têm desde há muito uma utilidade prática discutível, ou comprometida, tanto mais que se sabe que as normas dos testes de inteligência “enfraquecem” numa proporção de (pelo menos) 3 pontos por década (Flynn, 1984). Também por isso, as novas normas são preferíveis às mais antigas. A este propósito parece-nos importante retomar a posição de Kaufman (1993) quando afirma que as normas da

WISC-III são “excelentes” e “modernas”⁶. A possibilidade de recurso a normas nacionais recentes da WISC-III constitui uma vantagem decisiva para os psicólogos portugueses que fazem avaliação psicológica em crianças e adolescentes dos 6 aos 16 anos de idade.

Finalmente, é importante não descurar algumas limitações da presente investigação que impedem uma maior generalização dos resultados. O estudo aqui apresentado centra-se na comparação entre os desempenhos na WISC e na WISC-III, por parte de um grupo heterogéneo, do ponto de vista clínico, de crianças e adolescentes com dificuldades de aprendizagem e/ou problemas de comportamento. Ou seja, antes destes resultados poderem ser generalizados são necessários mais estudos, com grupos clínicos mais numerosos e homogéneos e, não menos importante, com sujeitos normais. Embora a direcção das discrepâncias entre os QIs obtidos a partir dos dois instrumentos seja inequívoca, admitimos que a sua magnitude possa ser diversa em grupos distintos do ponto de vista clínico ou educativo, bem como em diferentes idades e níveis de aptidão.

Previsivelmente, na sua prática profissional, os psicólogos portugueses irão substituir o uso da WISC pelo da WISC-III e necessitam, por isso, de mais

⁶ Num registo mais apologético e incisivo Kaufman (1994) escreve: “*far better than the outdated norms of the WISC-R: use the WISC-III*” (p. 38). Todavia, Sadlacek (1994) assinala que as amostras normativas mais recentes não são automaticamente amostras mais representativas e acrescenta que, em geral, os grupos minoritários se encontram insuficientemente representados nas referidas amostras normativas. Na opinião de Sadlacek esta seria uma outra hipótese explicativa para o declínio constantemente observado nos resultados.

informação específica relativa às comparações entre os resultados nestas duas versões. Coloca-se, pois, o problema da comparação de resultados. A implementação destas investigações é decisiva para os psicólogos que vão passar a usar na sua prática profissional a WISC-III (em detrimento da WISC). A informação daí derivada ajudará a clarificar o significado dos resultados desta pesquisa⁷ e a comparação dos desempenhos de sujeitos que responderam anteriormente à WISC e, a partir de agora, à WISC-III.

Quando é que devemos considerar que o QI de um sujeito se modificou? Se um sujeito foi anteriormente examinado através da WISC qual o seu resultado previsível na WISC-III, na hipótese de uma estabilidade das suas competências intelectuais? Para proceder a uma tal relação (o QI obtido na WISC e o intervalo no qual se situaria provavelmente o seu QI na WISC-III) e para interpretar mais compreensivamente estes resultados será necessária mais investigação, incluindo novos estudos com outras amostras constituídas quer por crianças e adolescentes sem problemas, quer por sujeitos com diagnósticos clínicos mais homogéneos.

Referências bibliográficas

Aichinger, K. S., Bolen, L. M., Hall, C. W., & Webster, R. E. (1993, April). *A comparison of the WISC-R and the WISC-III*. Paper presented at the annual meeting of the National Association of School Psychologists, Washington, DC.

⁷ Para além da importância das características dos sujeitos examinados, a variável examinador e as condições reais de aplicação das provas serão, nalguns casos, e apesar da natureza estandarizada das provas, aspectos igualmente a considerar aquando da comparação dos resultados.

- American Psychological Association (1999). *Standards for educational and psychological testing* (2nd ed.). Washington, DC: Author.
- Bolen, L. M., Aichinger, K. S., Hall, C. W., & Webster, R. E. (1995). A comparison of the performance of cognitively disabled children on the WISC-R and WISC-III. *Journal of Clinical Psychology, 51*, 89-94.
- Bryant, D. C. (1992). A comparison of WISC-R and WISC-III for gifted children. Unpublished master's thesis. West Virginia Graduate College, Morgantown, WV.
- Carlton, M. & Sapp, G. (1997). Comparison of WISC-R and WISC-III scores of urban exceptional students. *Psychological Reports, 80*, 755-760.
- Cassidy, L. C. (1998). The stability of WISC-III scores: For whom are triennial re-evaluations necessary? *Dissertation Abstracts International, Section B: The Sciences & Engineering*, Vol. 58 (8-B), pp. 4514.
- Dixon, W. E. & Anderson, T. (1995). Establishing covariance continuity between the WISC-R and the WISC-III. *Psychological Assessment, 7*, 115-117.
- Doll, B. & Boren, R. (1993). Performance of severely language-impaired students on the WISC-III, language scales, and academic achievement measures. *Journal of Psychoeducational Assessment [WISC-III Monograph]*, 77-86.
- Doppelt, J. E., & Kaufman, A. S. (1977). Estimation of the differences between WISC-R and WISC IQs. *Educational and Psychological Measurement, 37*, 417-424.
- Dumont, R. & Faro, C. (1993, May). The WISC-III: Almost two years old; proceeding with caution – practitioners' concerns. *NASP Communique, 21* (7), 12-15.
- Ferreira Marques, J. H. (1969). *Estudos sobre a Escala de Inteligência de Wechsler para Crianças (WISC): Sua adaptação e aferição para Portugal*. Dissertação de doutoramento em Filosofia apresentada à Faculdade de Letras da Universidade de Lisboa.
- Ferreira Marques, J. H. (1970). *Manual da Escala de Inteligência de Wechsler para Crianças (WISC)*. Lisboa: Instituto de Alta Cultura.
- Flieller, A. (2001). Problèmes et stratégies dans l'explication de l'effet Flynn. In M. Huteau (Dir.), *Les figures de l'intelligence* (pp. 43-66). Paris: Éditions et Applications Psychologiques.
- Flynn, J. R. (1984). The mean of IQ of Americans: massive gains 1932 to 1978. *Psychological Bulletin, 95*, 29-51.
- Flynn, J. R. (1985). Wechsler intelligence tests: do we really have a criterion of mental retardation? *American Journal of Mental Deficiency, 90*, 236-244.
- Flynn, J. R. (1987). Massive IQ gains in many nations: what IQ tests really measure. *Psychological Bulletin, 95*, 29-51.
- Flynn, J. R. (1998a). IQ gains over time: Toward finding the causes. In U. Neisser (Ed.), *The rising curve: Long-term gains in IQ and related measures* (pp. 25-66). Washington, DC: American Psychological Association.
- Flynn, J. R. (1998b). WAIS-III and WISC-III: US IQ gains 1972 to 1995: How to compensate for obsolete norms. *Perceptual and Motor Skills, 86*, 1231-1239.
- Flynn, J. R. (1999). Searching for justice: The discovery of IQ gains over time. *American Psychologist, 54*, 5-20.
- Flynn, J. R. (in press). The hidden history of IQ and special education: Can the problems be solved?
- Gaskill, F. W. & Brantley, J. C. (1996). Changes in ability and achievement

- scores over time: Implications for children classified as learning disabled. *Journal of Psychoeducational Assessment*, 14, 220-228.
- Graf, M. F. & Hinton, R. N. (1994). A 3-year comparison study of WISC-R and WISC-III IQ scores for a sample of special education students. *Educational and Psychological Measurement*, 54, 128-133.
- Grégoire, J. (1992). Comment interpreter la difference entre le QI Verbal et le QI de Performance au WISC-R. *Révue Européenne de Psychologie Appliquée*, 42, 29-33.
- Gridley, B. E., Arceneaux, J., Gayer, H., Buelow, B., & Palmer, C. (1994). Comparisons between WISC-R and WISC-III in reevaluation of special education students. Poster presented at the American Psychological Association Convention, Los Angeles, CA.
- Guilford, J. P., & Fruchter, B. (1978). *Fundamental statistics in psychology and education* (6th ed.). New York: McGraw-Hill
- Gunter, C. M., Sapp, G. L., & Greene, A. C. (1995). Comparison of scores on WISC-III and WISC-R of urban learning disabled students. *Psychological Reports*, 77, 473-474.
- Hamm, H., Wheeler, J., McCallum, S., Herrin, M., Hunter, D., Catoe, C. (1976). A comparison between the WISC and WISC-R among educable mentally retarded students. *Psychology in the Schools*, 13, 4-8.
- Harrison, P. L., Kaufman, A. S., Hickman, J. A., & Kaufman, N. L. (1988). A survey of tests used for adult assessment. *Journal of Psychoeducational Assessment*, 6, 188-198.
- Herring, S., & Reitan, R. M. (1986). Sex similarities in Verbal and Performance IQ deficits following unilateral cerebral lesions. *Journal of Consulting and Clinical Psychology*, 54, 537-541.
- Jones, P. E. (1996). Expected score differences in the WISC-R and WISC-III for special populations. *Dissertation Abstracts International, Section A: Humanities & Social Sciences*, Vol. 56 (10-A), pp. 4041.
- Kaufman, A. S. (1981). The WISC-R and learning disabilities assessment: State of the art. *Journal of Learning Disabilities*, 14, 520-526.
- Kaufman, A. S. (1990). *Assessing adolescent and adult intelligence*. Boston: Allyn & Bacon.
- Kaufman, A. S. (1993). King WISC the third assumes the throne. *Journal of School Psychology*, 31, 345-354.
- Kaufman, A. S. (1994). *Intelligent testing with the WISC-III*. New York: Wiley.
- Kaufman, A. S. & Lichtenberger, E. O. (2000). *Essentials of WISC-III and WPPSI-R assessment*. New York: Wiley.
- Klein, E. S., & Fisher, G. S. (1993, April). WISC-R and WISC-III: A comparison of IQ, subtest, and factor/index scores. Paper presented at the annual meeting of the National Association of School Psychologists, Washington, DC.
- Kogan, S. L. (1996). Comparison of the performance of L. D. children on the WISC-R and WISC-III. *Dissertation Abstracts International, Section A: Humanities & Social Sciences*, Vol. 59 (9-A), p. 3556.
- Larrabee, G. J., & Holyroyd, R. G. (1976). Comparison of WISC and WISC-R using a sample of highly intelligent children. *Psychological Reports*, 38, 1071-1074.
- Lyon, M. A. (1995). A comparison between WISC-III and WISC-R scores for learning disabilities reevaluations. *Journal of Learning Disabilities*, 28 (4), 253-255.
- Mackintosh, N. J. (1998). *IQ and human intelligence*. New York: Oxford University Press.

- Nagle, R. J. & Daley, C. E. (1994, March). *Longitudinal comparability of the WISC-III and the WISC-R with EMR students*. Paper presented at the annual meeting of the National Association of School Psychologists.
- Newby, R. F., Recht, D. R., Caldwell, J. & Schaefer, J. (1993). Comparison of WISC-III and WISC-R IQ changes over a 2-year time span in a sample of children with dyslexia. *Journal of Psychoeducational Assessment [WISC-III Monograph]*, 87-93.
- Post, K. R. (1992). *A comparison of WISC-R and WISC-III scores on urban special education students*. Unpublished educational specialists thesis, James Madison University, Harrisonburg, VA.
- Post, K. R. & Mitchell, H. R. (1993). The WISC-III: A reality check. *Journal of School Psychology*, 31, 541-545.
- Prifitera, A. & 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.
- Reynolds, C. R. (1990). Conceptual and technical problems in learning disability diagnosis. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children: Intelligence and achievement* (pp. 571-592). New York: Guilford.
- Sattler, J. M. (1992). *Assessment of children: WISC-III and WPPSI-R Supplement*. San Diego, CA: Jerome Sattler.
- Schwartz, F. G. (1976). A comparison of the WISC and WISC-R. *Psychology in the Schools*, 13, 1071-1074.
- Scott, M. A. (1999). A comparison of WISC-R and WISC-III IQ scores and special education placement in samples of Black and White mentally retarded and learning disabled students. *Dissertation Abstracts International, Section A: Humanities & Social Sciences*, Vol. 59 (9-A), pp. 3354.
- Sedlacek, W. E. (1994). Issues in advancing diversity through assessment. *Journal of Counseling & Development*, 72, 549-553.
- Sevier, R., & Bain, S. K. (1994). *Comparison of WISC-R and WISC-III scores for gifted students*. *Roeper Review*, 17, 39-42
- Simione, P. A. (1999). Comparative effects for learning disability assessment using the revised and third edition Wechsler Intelligence Scales for Children: Validity issues. *Dissertation Abstracts International, Section A: Humanities & Social Sciences*, Vol. 59 (9-A), pp. 3354.
- Sirois, P. A., Posner, M., Stehbins, J. A., Loveland, K. A., Nichols, S., Donfield, S. M., Bell, T. S., Hill, S. D., Amodei, N. & Hemophilia Growth and Development Study (2002). Quantifying practice effects in longitudinal research with the WISC-R and WAIS-R: A study of children and adolescents with hemophilia and male siblings without hemophilia. *Journal of Pediatric Psychology*, 27, 121-131.
- Slate, J. R. (1995). Two investigations of the validity of the WISC-III. *Psychological Reports*, 76, 299-306.
- Slate, J. R., Jones, C. H., & Saarnio, D. A. (1997). WISC-III IQ scores and special education diagnosis. *Journal of Psychology*, 131, 119-120
- Smith, D. K., Stovall, D. L., & Geraghty, B. L. (1994, March). *WISC-III/WISC-R relationships in special education re-evaluations*. Paper presented at the annual meeting of the National Association of School Psychologists, Seattle.

- Spafford, C. S. (1989). Wechsler Digit Span subtest: Diagnostic usefulness with dyslexic children. *Perceptual and Motor Skills*, 69, 115-125.
- Spitz, H. H. (1989). Variations in Wechsler interscale IQ disparities at different levels of IQ. *Intelligence*, 13, 157-167.
- Stavrou, E. (1990). The long-term stability of WISC-R scores in mildly retarded and learning-disabled children. *Psychology in the Schools*, 27, 101-110.
- Swerdlik, M. E. & Wilson, F. R. (1979). A comparison of WISC and WISC-R subtest scatter. *Journal of Learning Disabilities*, 12, 105-107.
- Vance, H., Maddux, C. D., Fuller, G. B. & Awadh, A. M. (1996). A longitudinal comparison of WISC-III and WISC-R scores of special education students. *Psychology in the Schools*, 33, 113-118.
- Vercoustre, O. (1985). WISC — WISC-R: Étude comparative des Q.I. *Revue de Psychologie Appliquée*, 35 (2), 105-124.
- Zimmerman, I. L. & Woo-Sam, J. M. (1996). Is retesting with the WISC-III a defensible procedure? *Perceptual and Motor Skills*, 82, 349-350.
- Webster, R. (1988). Statistical and individual temporal stability of the WISC-R for cognitively disabled adolescents. *Psychology in the Schools*, 25, 365-372.
- Wechsler, D. (1949). *Manual for the Wechsler Intelligence Scale for Children*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (1974). *Manual for the Wechsler Intelligence Scale for Children - Revised*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (1981). *WISC-R: Manuel*. Paris: Editions du Centre de Psychologie Appliquée.
- Wechsler, D. (1991). *Manual for the Wechsler Intelligence Scale for Children - Third Edition*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (1992). *WISC-III^{UK}*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (1996). *Manuel de l'Échelle d'Intelligence de Wechsler pour Enfants - Troisième Edition*. Paris: Editions du Centre de Psychologie Appliquée.
- Wechsler, D. (2003). *Escala de Inteligência de Wechsler para Crianças – Terceira Edição*. Lisboa: Cegoc.
- Weiner, S. G. & Kaufman, A. S. (1979). WISC-R versus WISC for black children suspected of learning or behavioral disorders. *Journal of Learning Disabilities*, 12, 100-105.