British Journal of Clinical Psychology (2005), 44, 279–288

The British Psychological Society
www.bpsjournals.co.uk

WAIS III UK: An extension of the UK comparability study

Robert Wycherley[1]*, Anthony Lavender[1], Sue Holttum[1], John R. Crawford[2] and Darren Mockler[3]
[1]Centre for Applied Social and Psychological Development, Canterbury Christ Church University College, UK
[2]School of Psychology, University of Aberdeen, UK
[3]King's Wood Centre, Colchester, UK

Objectives. To correct and augment the sample on which a previous study, assessing the applicability of USA WAIS III norms to the UK population, was based. To repeat this study using the modified sample.

Design. The WAIS III UK was administered to a supplementary sample, which was used to enhance and correct the original sample, allowing a better match to UK population demographics.

Method. WAIS III UK scores were processed using USA norms.

Results. Distributions of UK IQ, Index, and subtest scores were consistent with USA norms. Means were significantly above USA values and did not show a flat profile. Correlational analysis suggested the presence of an age-related sampling bias, which was then statistically controlled for using analysis of covariance, with age as covariate, to give corrected means. These means, with the exception of two subtests (block design and picture arrangement), did not differ significantly from USA values.

Conclusion. With a minor proviso concerning the above subtests, it is probable that USA norms can be safely used with the UK population. Future UK samples need to ensure random sampling within demographic categories if sampling bias is to be avoided.

The Wechsler adult intelligence scale III (WAIS III; Wechsler, 1997a) was published in the United States in 1997. Its predecessor, the WAIS – R (Wechsler, 1981), was Anglicized (Lea, 1986), but never standardized in the UK; the assumption being that the test would perform psychometrically as it did in the USA. A Scottish sample (Crawford, Gray, & Allan, 1995; Crawford & Allan, 1996) showed its psychometric properties to be similar to those of the USA normative sample, although intelligence quotient (IQ) means were a

*Correspondence should be addressed to Dr Robert Wycherley, Centre for Applied Social and Psychological Development, Salomons, Broomhill Road, Southborough, Tunbridge Wells, Kent TN3 0TG, UK (e-mail: b.j.wycherley@salomons.org.uk).

DOI:10.1348/014466505X29440

little higher, and their variance significantly less, than USA values. An earlier Scottish study (Crawford, Allan, Stephen, Parker, & Besson, 1989) found a factor structure closely matching that of the USA standardization sample. Overall, in these studies, the WAIS – R appeared to have similar psychometric properties in the UK and USA.

Prior to UK publication, the WAIS III was Anglicized and its psychometric performance with the UK population assessed (Wycherley, Benjamin, Crawford, & Mockler, 1999). Anglicization required no administration changes, and only minor changes to item content. Following Anglicization, a sample was collected which was intended to be representative of UK population for educational and occupational level, geographical area, gender, ethnic group, and age. Population data were derived from 1991 UK census data (OPCS, 1991: DOHSS NI, 1991), augmented by additional information from the Office for National Statistics, and the Department for Education and Employment. Participants were given the Anglicized WAIS III, and their scores were processed using USA norms. If these norms are applicable to the UK population, the sample's IQ and Index scores should be normally distributed, with means close to the USA mean of 100, and standard deviations close to 15. Similarly, subtest means and standard deviations should be close to 10 and 3, respectively.

Despite attempts to achieve a representative UK sample, this departed significantly from population expectations, and had to be matched to these by selectively eliminating participants. Sample and population frequencies were compared by means of chi-square tests, for each demographic variable. When a significant difference occurred, participants were randomly removed from the sample, until no significant difference remained. In this way, the original sample of 408 was reduced to a final sample of 332, which met expected population proportions for educational and occupational levels, gender and ethnic group, but departed significantly from these for geographic region and age.

IQ and Index distributions did not depart significantly from normality, but Verbal and Full-scale IQs, and the Verbal Comprehension Index, had standard deviations significantly below 15. Three subtest standard deviations were significantly below 3, and one significantly above.

IQ and Index means were close to, but (except for processing speed) significantly above 100. Six subtest means were also significantly above 10. The elevated means were interpreted in terms of a small degree of IQ drift (Flynn, 1998) and a degree of sampling bias, leading to the collection of a more able sample than had been intended. One feature that reflected this was the virtual absence of scores below three standard deviations from the mean. This had also occurred in the USA sample and was corrected by deliberately including participants with low IQ levels (Flynn, 1998). Overall, the results were similar to those of Crawford *et al.* (1995) for the WAIS – R, and suggested that, with minor provisions, the USA norms could reasonably be applied to the UK population.

Objectives

After publication of the WAIS III UK (Wechsler, 1999), limited data collection continued with the objective of remedying some of the deficiencies in the UK sample, and replicating the UK–USA comparability study with this modified sample. It was hoped that an augmented sample would allow demographic matching for geographical area and improve occupational matching. It was also hoped that a slightly enlarged sample might allow later UK–USA comparisons at a number of age-bands.

Method

Sampling procedure and data processing were identical to those of the main sample outlined earlier, and additional data amounted to 163 participants. As there had been a lack of scores in the lower tail of the original sample distribution, this was compensated for by including 11 clients of a learning disabilities service. The original sampling had collected data in six regions: Scotland, Northern Ireland, Wales, the Midlands, northern England and southern England. Additional data were collected in two previously under-represented areas, the Midlands and northern England, and attempts were also made to increase the representation of skilled, and partly skilled, manual and non-manual workers. All these categories were over-sampled to provide a pool of participants that could be drawn on to correct defects in the original sample. To allow later UK–USA comparisons at different age-bands, data collection tried to ensure reasonable numbers within age-bands rather than attempting to reproduce the age distribution of the population.

Demographic variables for the original UK sample were compared with expected population frequencies using chi-square tests. When a significant difference occurred, participants were added from the pool of new data, or removed from the original sample, until no significant differences remained. As far as possible, participants were added rather than removed, and where there was a choice of participant, this was made according to a table of random numbers. After re-matching there were no significant differences between the augmented sample and population proportions, for any of the demographic variables, and a virtually perfect match was achieved for education ($\chi^2 = 0.13, p = 1.00$). The final sample amounted to 370 participants, a modest increase of 38 over the original 332. Only eight participants with learning disabilities were included, so as not to exceed the expected population proportion of 2.2% of full scale IQ scores below 3 standard deviations from the mean. These participants were randomly selected from the 11 who had been sampled. All of the participant's WAIS III UK scores were processed according to USA norms.

As initial data analysis suggested that the augmented sample required statistical correction, a second analysis was undertaken. There are consequently two results sections, each followed by discussion.

Results I

The following analysis compares the augmented UK sample with USA norms. No statistical comparisons with the original UK sample will be made. In all cases of multiple comparison Bonferroni corrections were applied, and when evaluating repeated measures ANOVA effects, the conservative Greehouse–Geisser test was used when Mauchly's test of sphericity was significant.

Subtest standard deviations and means

Subtest standard deviations are shown in Table 1, and were compared with the USA value of 3 using the method described by Zar (1999). Only the standard deviation for *similarities* displayed a significant difference and was below the USA value ($p < .05$).

One sample *t* tests were used to compare each mean with the USA value of 10. From Table 1 it can be seen that 7 of the 14 means differed significantly from this value, all being above it. A repeated-measures ANOVA across subtests showed a significant subtest effect ($F = 11.075, p < .001$) indicating significant differences between subtest means.

Table 1. Subtest standard deviations and means

Subtest	*SD*	Mean
Picture completion	3.2	10.6[a]
Vocabulary	2.7	10.4
Digit symbol coding	3.0	10.0
Similarities	2.6[b]	10.0
Block design	3.1	11.3[a]
Arithmetic	3.1	10.5[b]
Matrix reasoning	3.0	10.9[a]
Digit span	3.1	10.9[a]
Information	2.7	11.1[a]
Picture arrangement	3.1	10.3
Comprehension	2.7	10.3
Symbol search	3.0	10.3
Letter-number sequencing	3.1	10.8[a]
Object assembly	3.0	10.3

[a] Significantly different from USA value of 10 with $p < .01$.
[b] Significantly different from USA value of 3 (*SD*) or 10 (mean) with $p < .05$.

IQ and Index distributions and means

Table 2 shows Kolmogorov–Smirnov test *Z* scores, which indicate that none of the IQ and Index distributions departed significantly from normality. Of the seven standard deviations only those for Verbal IQ and the Verbal Comprehension Index departed significantly from the expected USA value of 15 ($p < .05$, $p < .01$, respectively), both being below this figure.

Table 2. IQ and Index Kolmogorov–Smirnov Test *Z* scores, standard deviations, means, and Spearman correlation coefficients with age

IQ/Index	*Z* score	*SD*	Mean	*r* with Age
Full scale	1.3[a]	14.4	103.7[b]	0.15[c]
Verbal	1.1	13.4[d]	102.8[b]	0.11
Performance	0.9	15.8	104.3[b]	0.17[e]
Verbal comprehension	1.3	13.1[b]	102.4[b]	0.11
Perceptual organization	1.0	16.2	105.5[b]	0.15[c]
Working memory	1.4	15.8	104.2[b]	0.14[c]
Processing speed	1.4	14.6	100.9	0.04

[a] All *Z* scores non-significant.
[b] Significantly different from USA value of 15 (*SD*s) or 100 (means) with $p < .01$.
[c] Significant $p < .05$.
[d] Significantly different from USA value of 15 with $p < .05$.
[e] Significant $p < .01$.

Each IQ and Index mean was compared with the USA value of 100 using one-sample *t* tests. Table 2 shows that all except the Processing Speed Index were significantly above 100 ($p < .01$ in all cases).

Mean Performance IQ was significantly higher than mean Verbal IQ ($t = -2.380$, $p < .05$). Similarly, a repeated measures ANOVA for the four Indices showed a highly significant effect for Index ($F = 15.394$, $p < .001$). Additional *t* tests between Index pairs showed four of the six contrasts to be significant: Working memory was significantly higher than processing speed ($t = 4.256$, $p < .01$) and verbal comprehension ($t = -2.697$, $p < .05$), while perceptual organization was significantly higher than processing speed ($t = 6.029$, $p < .01$) and verbal comprehension ($t = -4.722$, $p < .01$). The most notable difference was that perceptual organization was significantly higher than verbal comprehension, just as Performance IQ had been significantly higher than Verbal IQ.

IQ and Index correlations with age

As each age-band of the USA sample is separately normed, significant correlations between age and intelligence measures for the overall UK sample (i.e. across all the age-bands) would not be expected, although small correlations within age-bands could arise. Table 2 shows positive UK correlation coefficients between IQ and Index scores and age, those for Full scale and Performance IQs, and Perceptual organization and Working memory Indices, reaching significance, though the effect sizes are small (Cohen, 1988, pp. 115–16). No significant correlation coefficients arose within any of the 11 age-bands after Bonferroni correction, indicating that the correlation between age, and the IQs and Indices, arose between, rather than within, the age-bands.

Discussion I

Only one of the subtests and two of the IQ and Index standard deviations differed significantly from the USA value, all three being below the appropriate figure. This indicates reasonable consistency between UK data and the USA norms.

Subtest means differed significantly from each other, and 7 of the 14 were significantly above the USA mean of 10. Similarly, with the exception of processing speed, IQ and Index means were significantly higher than USA means. Performance IQ was significantly higher than Verbal IQ, and means for perceptual organization and working memory were significantly higher than those for verbal comprehension and speed of processing. The IQ and Index measures thus showed two clusters: verbal and speed of processing measures being closest to the USA mean, while perceptual and working memory measures were significantly higher. This latter cluster, comprising Performance IQ, Perceptual Organization Index and Working Memory Index showed significant positive correlations with age (Table 2), whereas the first cluster (Verbal IQ, Verbal Comprehension Index and Processing Speed Index) did not. This indicates that the differences between these two clusters increased with age, although, as mentioned earlier, the effect was small.

There are a number of possible explanations for this pattern of elevated means, and their age-related differences. Four will be considered here: UK–USA physical differences, UK–USA cultural differences, IQ drift, and sampling bias.

UK–USA physical differences

It appears that the UK population, relative to that of the USA, shows a broad pattern of elevated subtest, IQ and Index means and a less steep decline in the Performance IQ, Perceptual Organization Index and Working Memory Index with age. A wide range of explanations for these differences could be advanced. One could hypothesize, for instance, that the elevated UK means were due to UK–USA genetic differences, or to environmental factors such as living conditions, nutrition or health care affecting cognitive functioning in different age cohorts, such that an apparently age-related effect is produced in the cross-sectional samples drawn upon here. It is probable, however, that such influences would produce rather broad cognitive effects, manifest across the range of sub-test, IQ and Index measures. It is more difficult to explain, in physical terms, the rather subtler USA–UK difference, at successive age-bands, of the Performance IQ, and the Perceptual Organization and Working Memory Indices. This is particularly so as other measures were not significantly affected, and any hypothesized influence would have to act on cohorts, not on individuals. For instance, one might argue that a larger proportion of older USA citizens had lived through poorer life conditions from childhood, for example, being born during the 1930s 'great depression'. Such circumstances could, theoretically, have had a differential effect on their cognitive development due to poor nutrition or other effects of poverty, in comparison to UK citizens born in the same period. The difficulty is that one would expect such a cohort effect to produce differences across all the subscales in the older cohorts, rather than only in some.

While physical explanations for elevated means cannot be ruled out, it is suggested that the specific age-related pattern found in the present project cannot be convincingly accounted for in this way. A telling piece of evidence comes from the study of the WAIS – R by Crawford *et al.* (1995) which produced overall UK means somewhat higher than those of the USA figures, but did not show the significant Verbal-Performance IQ difference that would have been expected if the apparent age-related pattern had been present at that time.

UK–USA cultural differences

Turning to cultural explanations, considerable efforts were made during WAIS III development to eliminate culturally biased items (Wechsler, 1997b, pp. 10–11), making differences attributable to the relative proportions of ethnic minority groups less probable. However, if such differences did occur, they would be expected to show a fairly consistent effect across IQ and Index measures (Kaufman, 1990, pp. 158–163). As the proportion of ethnic minority groups in the USA is greater than in the UK, any such difference incorporated in the USA norms, would give a small advantage to the UK population, producing slightly higher mean scores. What would not be predicted is the pattern of age-related discrepancies that emerged from the present sample. Again, the previously cited WAIS – R study of Crawford *et al.* (1995), while supporting the possibility that the UK might have a small cognitive superiority, showed no evidence of the present pattern of means. Overall, physical or cultural explanations could reasonably account for a degree of general UK cognitive superiority, but they do not adequately explain the present age-related pattern of differences.

IQ drift

Flynn (1998) noted a rise in IQ measures over time and estimated this at about 0.25 of a point per year overall, though the figure for Performance IQ was greater than for Verbal IQ. The mid-point for the collection of the USA sample was 1996, and for the present sample 1998, giving an estimated rise for verbal measures of about 0.2 and, for non-verbal measures, about 0.6 of a point. Clearly, IQ drift, while consistent with the pattern of IQ and Index differences found in the present study, cannot explain its magnitude, or its increase with age.

Sampling bias

The fourth possible explanation is sampling bias, particularly as the present study used volunteer participants. Rosenthal and Rosnow (1975) suggest that volunteers tend to be significantly more intelligent than non-volunteers and estimate (pp. 68) a median inflationary effect of 0.18 of a standard deviation for IQ measures. In the present sample this would have amounted to 2.7 points for IQs and Indices. Coupled with the estimated IQ drift suggested in the preceding paragraph, correction for such bias would eliminate any substantial difference between UK and USA means, and some would fall below 100.

While volunteer bias and IQ drift plausibly explain the magnitude of the present UK and USA differences in means, the discrepancies between the two IQ and Index 'clusters', and their increase with age, also has to be accounted for. This pattern may be explained by the fact that, in line with population statistics, the sample's overall educational level declined with age, particularly as increasingly higher proportions of participants fell into the category 'no educational qualifications' as age increased. The proportion of the population falling into this category rises from below 10% at age 18–19 to over 60% at age 75–79, and it probably embraces a wide range of intellectual ability, including those whose actual ability exceeded their educational attainments because of lack of educational opportunity. Coupled with volunteer bias, the progressive expansion of this category may have allowed an increasing proportion of intellectually able, but poorly educated, participants to enter the sample. As educational level correlates more highly with verbal intelligence than with non-verbal intelligence (Kaufman, 1990, pp. 174–176) able, but poorly educated participants would probably show a pattern in which their Verbal IQ and Verbal Comprehension Index was lower than their Performance IQ and Perceptual Organization Index.

Before completing this discussion of sampling bias, it should be noted that the USA norms were also based on a volunteer sample, which was potentially open to biases similar to those described above. However, details of the collection of this sample (Wechsler 1997b, p. 21) suggest that recruitment of participants within its demographic categories was more probably random than was the case for the UK sample. Volunteer bias in the USA sample would, of course, probably lead to over-recruitment of more able participants, and to norms that were more stringent than they should be, producing UK means that were below USA values. Thus, if volunteer bias was present in the USA sample, the present pattern of UK means would indicate a greater degree of UK bias than has been presumed.

To summarize, it has been argued that the age-related pattern of discrepancies found in the present sample was probably due to a modest element of IQ drift and a larger degree of volunteer bias. This, coupled with an age-related weakening of educational level within the sample, led to the recruitment of an increasing proportion of participants whose verbal intellectual level fell below their non-verbal level.

In relation to this pattern, the Processing Speed Index mean is anomalous in its lack of elevation, and its lack of correlation with age (Table 2) shows it to be impervious to the postulated volunteer bias. Leaving aside statistical significance, if the correlations with age shown in Table 2 are taken to reflect the degree of volunteer bias affecting each IQ and Index, these would suggest that non-verbal measures were the most affected, verbal measures were less affected, and processing speed was hardly affected at all. Seen in this light, it would represent the true 'unbiased' intellectual level of the sample, to which the other measures would tend in the absence of age-related volunteer bias.

Results II

It was decided to test the explanation advanced in the previous section by statistically controlling for the correlation between IQ, Index, and subtest scores, and age (i.e. the presumed sampling bias), using analyses of covariance (ANCOVA), with age as the covariate.

Separate ANCOVAs for the Verbal and Performance IQs, and for the Indices, produced no significant IQ or Index effects, but there were significant interactions between IQ and age ($F = 4.512$, $p < .05$), and between Index and age ($F = 2.643$, $p = .05$). These results indicated that the former significant differences between Verbal and Performance IQ, and between the four Indices, disappeared when their correlation with age was controlled for.

ANCOVA for subtests showed a significant effect for subtest, and a significant age-subtest interaction ($F = 3.843$, $F = 3.933$, $p < .01$ in both cases). 'Adjusted' subtest means, for which the presumed age-related bias had been controlled, together with their 95% confidence intervals, were derived from this ANCOVA. Confidence intervals for block design, with a mean above the USA value of 10, and picture arrangement, with a mean below this value, failed to span the USA mean, suggesting that the significant subtest effect was attributable to them. Confidence intervals have been used here in place of more familiar tests, with a significant difference with $p < .05$ being indicated if the USA value falls outside a given 95% confidence interval. Removal of both these subtests from the analysis removed the significant subtest effect, leaving only a strongly significant age-subtest interaction ($F = 2.978$, $p < .01$). This showed that the remaining means did not differ significantly from each other when their correlation with age was controlled for.

Table 3 shows the adjusted means for IQs, Indices and subtests derived from the above analyses, with their 95% confidence intervals. Confidence intervals for all the IQ and Index means spanned the USA value of 100, indicating that none of these measures differed significantly from this figure, and all but one of the means were slightly below this value. Similarly, with the exception of the two subtests mentioned above, none of the subtest means departed significantly from the USA value of 10, and they showed small variations either side of this value.

Discussion II

The adjusted means shown in Table 3 were consistent with the argument advanced in Discussion I that most of the difference between USA and UK means could be attributed to age-related sampling bias. If this argument is accepted, then UK IQ and Index means appear to be marginally, though not significantly, below 100. Even allowing for a degree of IQ drift, the differences in UK and USA means would not be statistically significant, and would be clinically insignificant. With the exception of block design and picture arrangement, the adjusted subtest means were clustered around the USA mean of 10, and did not differ significantly from this value.

Block design and picture arrangement are of interest as possibly reflecting highly specific USA–UK cultural differences though, clinically, their deviation would be of little significance. As they deviate in opposite directions they would cancel out each other's effects within the Performance IQ; though, as picture arrangement does not contribute to Index scores, the Perceptual Organization Index would be marginally raised by block design, though not to a clinically significant degree. The greater difficulty of picture arrangement for the UK population could be due to lack of familiarity with some of

Table 3. 'Adjusted' IQ, Index and subtest means, with 95% confidence intervals

IQ/Index	Mean and CI
Full scale	99.0 (95.0, 103.0)[a]
Verbal	99.7 (96.4, 102.9)
Performance	98.4 (94.6, 102.1)
Verbal comprehension	99.4 (96.2, 102.5)
Perceptual organization	100.4 (96.5, 104.3)
Working memory	99.4 (95.6, 103.2)
Processing speed	99.7 (96.1, 103.2)
Subtest	Mean and CI
Picture completion	9.5 (8.7, 10.3)
Vocabulary	9.9 (9.3, 10.6)
Digit symbol coding	9.7 (9.0, 10.5)
Similarities	9.6 (9.0, 10.3)
Block design	11.2 (10.5, 12.0)[b]
Arithmetic	9.9 (9.1, 10.6)
Matrix reasoning	9.6 (8.9, 10.3)
Digit span	9.9 (9.2, 10.6)
Information	10.2 (9.6, 10.9)
Picture arrangement	8.8 (8.1, 9.6)[b]
Comprehension	10.3 (9.6, 10.9)
Symbol search	10.2 (9.5, 10.9)
Letter-number sequencing	10.0 (9.3, 10.8)
Object assembly	10.2 (9.5, 10.9)

[a] Estimated from sums of scaled scores for verbal and performance IQ.
[b] Confidence intervals that do not include 100 (IQ/Index means), or 10 (subtest means), indicating a significant UK–USA difference with $p < .05$.

the depicted scenes. Pizza tossing, surfboards, and sharks are probably less familiar to the UK than to the USA population. Turning to block design, Kaufman (1990, p. 161) indicates that the WAIS – R block design subtest showed a strong black–white racial difference in favour of whites and, given the difference in ethnic mix of the two populations, a small UK advantage could be built into the norms. However, these two subtests were part of the WAIS – R, and their means were not unduly discrepant in the only UK study of this test (Crawford *et al.*, 1995). Confirmation that they are indeed discrepant must await independent sampling, though, in the mean time, it would be clinically wise to interpret their results with care.

If it is accepted that the pattern of unadjusted means found in the present study was predominantly due to age-related sampling bias, then future research should attempt to overcome this by ensuring truly random sampling within demographic categories. This is particularly necessary for the major demographic variable of educational level, and especially when a large proportion of participants is covered by a single demographic category. This arises most crucially in the older age-bands, for which high proportions of participants are contained in the categories 'no educational qualifications' and 'occupationally inactive'. A similar problem arises with the youngest age-bands as 'degree' and 'diploma' categories are all but absent for the education variable, reducing the number of educational categories. Many younger participants also fall in the 'occupationally inactive' category as they are students. Such categories potentially contain wide ranges of ability, and sampling bias is likely to become a problem if volunteers are used, and sampling is not truly random.

Another implication is that future research should relate to specific age bands rather than combining these into a single sample that might meet overall demographic criteria but have mis-matches for particular age-bands. As the WAIS III is separately normed for each age-band there is no reason why research into a single age band, or a small group of age bands, should not be undertaken. UK research into the two oldest age bands in the USA sample, 80–84, and 85–89 is particularly needed, as these were not covered by the present UK sample.

Acknowledgements

The authors would like to thank Mr. Paul McKeown, Director of the Psychological Corporation, Europe, for allowing use of the original WAIS III UK comparability sample and Dr Nigel Beail of Barnsley Community and Priority Services NHS Trust, for collecting the learning disabilities sample.

References

Crawford, J. R., & Allan, K. M. (1996). WAIS – R subtest scatter: Base-rate data from a healthy UK sample. *British Journal of Clinical Psychology*, *35*, 235–247.

Crawford, J. R., Allan, K. M., Stephen, D. W., Parker, D. M., & Besson, J. A. O. (1989). The Wechsler adult intelligence scale – revised (WAIS – R): Factor structure in a UK sample. *Personality and Individual Differences*, *10*, 1209–1212.

Crawford, J. R., Gray, C. D., & Allan, K. M. (1995). The WAIS – R (UK): Basic psychometric properties in an adult UK sample. *British Journal of Clinical Psychology*, *34*, 237–250.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). New Jersey: Lawrence Erlbaum Associates.

Department of Health and Social Services and Registrar General for Northern Ireland. *The Northern Ireland Census 1991*. Belfast: HMSO.

Flynn, J. R. (1998). WAIS III and WISC III IQ gains in the United States from 1972–1995: How to compensate for obsolete norms. *Perceptual and Motor Skills*, *86*, 1231–1239.

Kaufman, A. S. (1990). *Assessing adolescent and adult intelligence*. Boston: Allyn and Bacon.

Lea, M. (1986). *WAIS – R. UK. A British supplement to the manual of the Wechsler adult intelligence scale – revised*. London: The Psychological Corporation.

Office of Population Censuses and Surveys. (1991). *Census Report for Great Britain*. London: HMSO.

Rosenthall, R., & Rosnow, R. L. (1975). *The volunteer subject*. New York: Wiley.

Wechsler, D. (1981). *Wechsler adult intelligence scale – revised*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997a). *Wechsler adult intelligence scale* (3rd ed). San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997b). *WAIS-III, WMS-III technical manual*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1999). *Wechsler adult intelligence scale-third edition-UK*. London: The Psychological Corporation.

Wycherley, R. J., Benjamin, L., Crawford, J., & Mockler, D. (1999). Description of the UK study. In D. Wechsler (Ed.), *Wechsler adult intelligence scale-third edition-UK. Administration and scoring manual* (pp. 215–233). London: The Psychological Corporation.

Zar, J. H. (1999). *Biostatistical analysis*. Upper Saddle River, NJ: Prentice Hall.

Received 30 December 2002; revised version received 10 March 2004